Population connectivity of the Southern rock lobster, *Jasus edwardsii*

by

Cecilia Villacorta-Rath

BSc, MSc

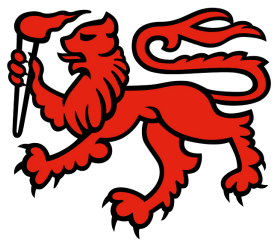
Institute for Marine and Antarctic Studies (IMAS)

Submitted in fulfillment of the requirements for the degree of

Doctor of Philosophy

August 2018

University of Tasmania



UNIVERSITY*of*
TASMANIA

IMAS
INSTITUTE FOR MARINE
& ANTARCTIC STUDIES

STATEMENTS AND DECLARATIONS

Declaration of Originality

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

Authority of Access

This thesis may be made available for loan and limited copying and communication in accordance with the Copyright Act 1968.

Statement regarding published work contained in thesis

The publishers of the papers comprising chapters 2 and 3 hold the copyright for that content, and access to the material should be sought from the journal. The remaining non-published content of the thesis may be made available for loan and limited copying and communication in accordance with the Copyright Act 1968.

Signed

Cecilia Villacorta-Rath
Date: 09/08/2018

STATEMENT OF CO-AUTHORSHIP

The following people and institutions contributed to the publication work undertaken as part of this thesis:

- Cecilia Villacorta-Rath, *IMAS, University of Tasmania* = **Candidate**
- Irina Ilyushkina, *School of Biological Sciences, Victoria University of Wellington* = **Author 1**
- Jan M. Strugnell, *Department of Ecology, Environment and Evolution, School of Life Sciences, La Trobe University* = **Author 2**
- Bridget S. Green, *IMAS, University of Tasmania* = **Author 3**
- Nicholas P. Murphy, *Department of Ecology, Environment and Evolution, School of Life Sciences, La Trobe University* = **Author 4**
- Stephen R. Doyle, *Department of Animal, Plant and Soil Sciences, School of Life Sciences, La Trobe University* = **Author 5**
- Nathan E. Hall, *Genomics Research Platform, La Trobe University* & Life Sciences Computation Centre, Victorian Life Sciences Computation Initiative = **Author 6**
- Andrew J. Robinson, *Genomics Research Platform, La Trobe University* & Life Sciences Computation Centre, Victorian Life Sciences Computation Initiative = **Author 7**
- James J. Bell, *School of Biological Sciences, Victoria University of Wellington* = **Author 8**
- Carla A. Souza, *Department of Ecology, Environment and Evolution, School of Life Sciences, La Trobe University* = **Author 9**

- Caleb Gardner, *IMAS, University of Tasmania* = **Author 10**

Author details and their roles in published paper:

Outlier SNPs enable food traceability of the southern rock lobster, *Jasus edwardsii*. Located in chapter 2. This work has been published in Marine Biology (2016) 163: 223. Cecilia Villacorta-Rath, Irina Ilyushkina, Jan M. Strugnell, Bridget S. Green, Nicholas P. Murphy, Stephen R. Doyle, Nathan E. Hall, Andrew J. Robinson, James J. Bell.

Candidate was the primary author and was largely responsible for the data analysis and paper writing, Author 1 performed laboratory work, assisted with data analysis, contributed to the idea, its formalization and manuscript preparation. Author 2, Author 3 and Author 8 acted as supervisors for this publication and contributed intellectually to the study design and manuscript preparation. Author 4 contributed intellectually to manuscript preparation. Author 5 contributed with laboratory work, data analysis and manuscript preparation. Author 6 and Author 7 provided bio-informatics support.

Temporal genetic patterns of diversity and structure evidence chaotic genetic patchiness in a spiny lobster. Located in chapter 3. This work is currently *in press* for publication in Molecular Ecology. doi: 10.1111/mec.14427. Cecilia Villacorta-Rath, Carla A. Souza, Nicholas P. Murphy, Bridget S. Green, Caleb Gardner, Jan M. Strugnell.

Candidate was the primary author and was largely responsible for the data analysis and paper writing. Author 2, Author 3 and Author 10 acted as supervisors for this publication and contributed intellectually to the study design and manuscript preparation. Author 4 contributed intellectually to manuscript preparation. Author 9

performed laboratory work, assisted with data analysis and manuscript preparation.

Can dispersal history drive spatio-temporal phenotypic differences between Southern rock lobster (*Jasus edwardsii*) post-larvae? Located in chapter 4. This work is in preparation for publication. Candidate was the primary author and was largely responsible for the data analysis and paper writing. Author 2, Author 3 and Author 10 acted as supervisors for this publication and contributed intellectually to the study design and manuscript preparation. Author 4 contributed intellectually to manuscript preparation. Author 9 performed laboratory work, assisted with data analysis and manuscript preparation.

Preliminary assignment of *Jasus edwardsii* post-larvae to population of origin. Located in chapter 5. Candidate was the primary author and was largely responsible for the data analysis and paper writing. Author 2, Author 3 and Author 10 acted as supervisors for this publication and contributed intellectually to the study design and manuscript preparation.

Signed

Cecilia Villacorta-Rath
PhD Candidate
Institute for Marine and Antarctic Studies – Fisheries and Coasts Centre
University of Tasmania

Signed

Associate Professor Bridget S. Green
Primary Supervisor
Institute for Marine and Antarctic Studies – Fisheries and Coasts Centre
University of Tasmania

Signed

Professor Chris Carter
Director Centre Head Fisheries and Aquaculture Centre
Institute for Marine and Antarctic Studies
University of Tasmania

ACKNOWLEDGEMENTS

This thesis would not have been possible without the help of fishermen, technical officers and fellow students who contributed to sample collection. Thanks to Andrew Kent, owner of the "Shelle-ton"; Peter Warner, owner of the "Livingstone"; Ben Lowler, owner of the "Tru Blu"; and Kent Way, owner of the "Chieftain G" for collection of adult lobster samples on the east and west coasts of Tasmania. Also thanks to Gary Carlos and Colin Fry for field assistance, fieldtrip coordination and logistics and adult lobster sample collection. Thanks to Felipe Briceño for field assistance in adult lobster collection. Thanks to Peter Hawthorne (SARDI) for adult and puerulus lobster collections in South Australia and Daniel Ierodiaconou (Deakin University) for adult collection in Victoria. Thanks to Ivan Hinojosa, Ruari Colquhoun, David Fallon, Kylie Cahill, Graeme Ewing, Sarah Pyke for puerulus lobsters collection in Tasmania.

I would like to express my deepest gratitude and admiration to Associate Professor Bridget Green and Associate Professor Jan Strugnell who encouraged and supported me along this journey. I considered it an honor to work with such bright and generous women. I would like to acknowledge Professor Caleb Gardner for enhancing this thesis with spot on comments. I am very grateful to Dr. Nick Murphy for his significant input developing my chapter manuscripts and co-authoring the papers coming out of this thesis. Special thanks to Dr. Karen Miller who kicked-off this project.

I am indebted to Dr. Stephen R. Doyle who patiently introduced me to the world of next generation sequencing and guided me through my first ddRADseq libraries

without losing the smile on his face. Special thanks for bioinformatics and statistics support to Dr. Nathan E. Hall, Andrew J. Robinson, Dr. Ira Cooke and Dr. Simon Watson. Thanks to Carla Souza for teaching me good laboratory practices and helping with bioinformatics.

Thanks to my many fellow PhD students for supporting me during hard times of laboratory work, data analysis and thesis write-up: Irina Ilyushkina, Laura N. Woodings, Michael D. Amor, Erin Hill, Nikeisha Caruana, Brett P. Shiel, Sussana Brycesson, Jude Hatley, Cecilia Carrea, Leslie-Ann Watson and Richard Lamedin. Thanks to Mel Best, Brooke Whitelaw and Daniel Laird for genetics laboratory assistance, and to Colin Fry, Anaïs Camberabero, Marilia Erickson and Madison Sankovitz for helping with puerulus weighing. Thanks to the Molecular Ecology Lab post-docs, Álvaro Roura and Katherine Harrisson for constructive conversations. Thanks to Adam Smolenski, Facility Manager of the Molecular Genetics Facility, and Lisette Robertson, Laboratory Coordinator, Health and Safety Representative. Big thanks to Lynne Baily and Gail Richardson for facilitating tedious admin work, fixing my mistakes and help getting last-minute signatures so many times.

I would also like to thank Hillary Revill (DPIPWE) and the Tasmanian Rock Lobster Fisherman Association (TRLFA) for their eagerness to know more about the complicated dispersal patterns of these amazing creatures. Thanks to the Institute for Marine and Antarctic Studies (IMAS) and the Graduate Research Office from the University of Tasmania for their financial support to attend international conferences.

Thanks to all the talented scientists who I can call friends, for the good times and interesting conversations: Tania Mendo, Kelly Ortega, Sam Noonan, Miwa

Takahashi, Anita Giraldo, Zouzou, Vero Mocellin, Daniela Guevara, Kor-jent vanDijk, Marios Aristophanous, Kyvely Tziouveli, James White and Bridie Allan. Respect to you all.

Thanks to my parents for encouraging me to pursue a degree and for always supporting me from far away.

Finally, thanks to my beautiful partner, Dimitri. Without your immense support and understanding I would not have been able to achieve this. Thanks for having food always ready when I needed it the most and a warm *pain au chocolat* waiting for me after working crazy long hours in the lab. Thanks for being so flexible and putting up with moving cities so many times. Thanks for tolerating my crankiness when things in the lab didn't work out. Thanks for the beautiful family that we now have.

To Yitzhak and
Dimitri,

GENERAL ABSTRACT

Understanding the mechanisms driving larval dispersal and connectivity is of ecological relevance and is beneficial in fisheries management. Defining population structure and stock boundaries helps in management of spawning stock biomass and annual harvests. Moreover, determining if there are patterns in dispersal can help identifying source populations that need management strategies directed towards maintaining appropriate levels of egg production.

The Southern rock lobster, *Jasus edwardsii*, extends around southeast Australia and New Zealand and supports valuable fisheries in both countries. Adults *J. edwardsii* do not migrate and their phyllosoma larvae, the dispersal stage, are adapted for drifting for approximately 12 to 24 months of pelagic larval duration. Consequently there has been an assumption of genetic homogeneity within the population throughout Australia. The assumption of panmixia has been supported by larval transport simulations and previous studies on genetic connectivity. The general eastward flow of currents in southeast Australia has been identified as the likely main dispersal mechanism. South Australia is a highly productive area, and is predicted by oceanographic models to be a source of larvae to the Tasmanian fishery. A second prediction from larval transport simulations is that regional self-recruitment varies markedly across the species range in Australia. Long-term monitoring of recruitment throughout the fishery shows high year-to-year variability in recruitment, as well as regional fluctuations. This has been linked to changes in environmental conditions. Fluctuations in recruitment magnitude can reduce the accuracy of population modeling of the stock, which is used to determine harvest strategies.

In this thesis I assessed variability in genetic identity of Southern rock lobster at different spatio-temporal scales to evaluate drivers of population structure. I reviewed

possible biological, environmental (e.g., dispersal history) or adaptive drivers (e.g., natural selection) by analyzing single nucleotide polymorphisms (SNPs) markers generated using double digest restriction site-associated DNA sequencing (ddRADseq). I measured genetic variability in *J. edwardsii* across three spatial scales, broad (1,000's km), medium (100's km) and fine-scale (10's km), as well as two temporal scales, within a year and between years.

In chapter 2, large-scale connectivity and potential for local adaptation between adult *J. edwardsii* from Australia and New Zealand was investigated using neutral and outlier markers. There was large-scale genetic divergence between Australia and New Zealand, two countries thousands of kilometres apart, at neutral regions of the genome (F_{ST} = 0.022), supporting previous findings of limited larval dispersal across the Tasman Sea. A much larger genetic differentiation was detected (F_{ST} = 0.134), using regions of the genome under putative selection suggesting local adaptation and post-settlement mortality of unfit genotypes.

In chapter 3 I assessed the extent and patterns of genetic variability in new recruits through time on a medium spatial scale. To determine the role of genetics in the observed interannual variability and how post-settlement selection acts to modify the observed structure in recruits, pueruli and post-pueruli settling during four consecutive years were analyzed. Interannual genetic variability of recruits within and between two sites located 100's of kilometres apart in South Australia and Tasmania provided support for chaotic genetic patchiness. Lower genetic diversity was observed during years of low puerulus catch rates at the Tasmanian site, suggesting regional genetic differences in recruitment. Additionally, the magnitude and strength of genetic divergence detected in the markers under putative positive selection also exhibited temporal and spatial variability. Both locations exhibited a single marker

under putative positive selection in common across years, providing weak evidence for post-settlement selection.

In chapter 4 I assessed fine-scale temporal and spatial genetic and phenotypic divergence in recruits across a latitudinal gradient. This was done using new recruits within one recruitment season in Tasmania from sites 10's of kilometres apart. There was a lack of overall population structure identified between three sites along a latitudinal gradient, but genetic divergence at a small spatial scale suggested chaotic genetic patchiness. Individuals sampled from the southernmost site during three consecutive monthly collections were genetically divergent from each other. There were also phenotypic differences of pueruli between sites and months of settlement; individuals at the northernmost site were consistently smaller at settlement. Collective dispersal is a possible mechanism of larval *J. edwardsii,* based on significant phenotypic differences between sites that were persistent through time. This implied that larvae released during the same spawning event could maintain cohesiveness until settlement, leading to genetic patchiness among individuals recruiting during the same year.

In chapter 5 I tested a number of the dispersal pathways projected by a larval transport simulation model for *J. edwardsii*. The predicted population of origin of pueruli caught in collectors in South Australia and Tasmania was then tested using genetic assignment to determine if they were the likely point of origin. All three adult lobster sampling sites were assigned as equally likely source for recruits in both South Australia and Tasmania. These results further evidenced the high level of genetic exchange in the Australian *J. edwardsii* population.

In general, the findings of this thesis provide new evidence on the dispersal mechanisms used by larval *J. edwardsii* driving the observed genetic variation in recruits. The high level of genetic admixture found herein highlights the need of a

coordinated fisheries management strategy between states in order to protect sub-populations that constitute important source of recruits.

TABLE OF CONTENTS

Chapter 1: General Introduction

Knowledge of connectivity in marine organisms is not only of ecological relevance but can also inform fisheries management. Fisheries management units are commonly based on jurisdictional limits, which do not necessarily match biological boundaries (Bernatchez et al. 2017). Identification of populations contributing to the next generation of recruits is of prime importance for fisheries management. Because larval dispersal in the marine environment is frequently unequal, some regions along a species geographic range are more vulnerable to overfishing (Fu and Fanning 2004; Truelove et al. 2015). Refining management plans based on biological populations and considering the directionality of the migration can help manage a stock's sustainable yield more effectively (Kerr et al. 2014).

Even in the presence of the high level of genetic admixture that characterizes most marine populations, selective pressure can cause local adaptation, resulting in cryptic population differentiation (Nielsen et al. 2009a; Corander et al. 2013). Disentangling the adaptive selection versus neutral processes as well as the extent of the influence of the environment on selection can also aid fisheries management (Nielsen et al. 2009a). This is because while larval transport simulations can estimate migration rates, but they usually do not take into account selective mortality. Therefore, an understanding of how connectivity is shaped by larval dispersal and natural selection is needed.

1.1 Factors affecting larval dispersal

A historical paradigm in population genetics is that a short pelagic larval duration (PLD) makes larvae more prone to settling close to their parents, generating genetic structure in the population (Palumbi 1994). On the contrary, long PLDs lead to long-

distance dispersal and therefore population admixture (Bohonak 1999). However, recent evidence suggests that population structure can exist in species with protracted PLDs (Shanks et al. 2003; Teske et al. 2015). The indication of these contrasting patterns is that it is not possible to draw generalizations on the relationship between duration of the larval phase and the resulting connectivity. Rather, there are a number of abiotic and biotic variables that play an important role influencing the dispersal kernel (Shanks 2009).

Among the biotic factors, larval behavior can affect the retention or advection of individuals into particular areas. Fish and invertebrate larvae can alter their position in the water column in order to forage and avoid predation (Sponaugle et al. 2002; Butler IV et al. 2011). This behavior can interact with hydrodynamic features and hinder dispersal over long distances by retaining larvae in coastal zones (Marta-Almeida et al. 2006; Butler IV et al. 2011). Biotic barriers, such as changes in productivity and presence of predators, can also prevent long distance dispersal of larvae and create genetic subdivision (Sponaugle et al. 2002). Among the abiotic factors influencing larval dispersal, hydrodynamic features can favor dispersal even in the presence of short PLDs. For example, barnacles with a PLD of two weeks disperse over hundreds of kilometres due to physical transport in the intertidal area, leading to low population structure (Gaines and Roughgarden 1985). Understanding the processes that could lead to larval dispersal in each species is necessary before reaching generalizations about genetic connectivity based on the length of the larval phase (Shanks 2009).

1.2 Population structure in recruits

In many marine invertebrates, newly settled organisms exhibit heterogeneous distribution of genotypes (Johnson and Black 1982; Johnson and Wernham 1999;

Kennington et al. 2013a; Truelove et al. 2017), which can persist into the adult population (Iacchei et al. 2013). This genetic heterogeneity, termed "chaotic genetic patchiness", is due to the interplay of selective and stochastic events. Natural selection can exert pressure on larvae and remove unfit genotypes from the pool of recruits (Johnson and Black 1992). However, the selective pressure can vary from one recruitment event to the other due to environmental stochasticity (Johnson and Black 1992). Moreover, variability in the parental stock can also produce genetic divergence among recruits (Eldon et al. 2016). Given the multifactorial nature of larval dispersal, the resultant genetic heterogeneity is therefore ephemeral (Johnson and Black 1992).

Variability in the parental stock can produce a sweepstakes-like scenario whereby only a few "winners" replenish the future populations. Generally it occurs in species with high female fecundity, high dispersal potential and low to moderate levels of population structure (Je Lee and Boulding 2009). The effect of environmental stochasticity on larval mortality would result in the reproductive success of a small minority of individuals (Hedgecock and Pudovkin 2011). Diversifying selection taking place prior and post-settlement can also result in a patchy distribution of genotypes (Eldon et al. 2016). Selection can produce genetic patchiness because it not only varies in strength among years but also from micro geographic scales up to broad latitudinal differences (Johnson and Black 1982; Johannesson et al. 1995; Veliz et al. 2006). Finally, collective dispersal can produce genetic patchiness in settlers due to the simultaneous arrival of cohorts of individuals released during the same spawning event (Funes-Rodríguez et al. 2015).

1.3 Implications of larval dispersal and connectivity in a commercial species

The commercially valuable Southern rock lobster, *Jasus edwardsii*, is a stuitable species to examine connectivity and larval dispersal at different scales. This species

has a broad geographical distribution, extending from the southern part of Western Australia through to Tasmania and New Zealand (Fig. 1.1). It supports an important fishery, providing revenue of US$292 per year into Australia (ABARE-BRS 2009). *J. edwardsii* is a long-lived species, with a life span of more than 30 years (Godwin et al. 2011) and adults reach sexual maturity between two and seven years of age, depending on the latitude (Annala et al. 1980).

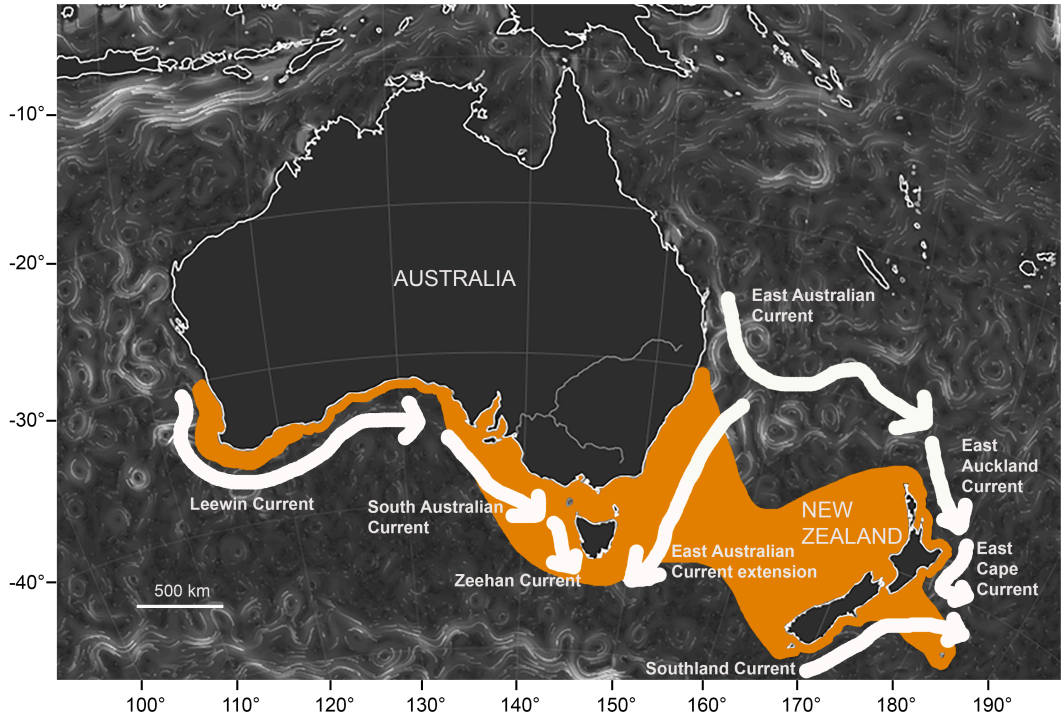


Fig. 1.1 Geographic distribution of *Jasus edwardsii* (highlighted in orange). White arrows indicate the prevailing ocean currents in Australia and New Zealand. Background image from: earth.nullschool.net

As many other invertebrates, *J. edwardsii* is highly fecund, each female can produce up to half a million eggs per year (Green et al. 2009). They reproduce once a year, incubating eggs from May to September, but egg-hatching periods vary between sub-populations across a latitudinal gradient (Powell et al. 2016). Phyllosoma larvae move to the water surface shortly after hatching (Booth and Phillips 1994) and their

flat bodies enable them to drift passively in the ocean (Phillips and McWilliam 1986). Phyllosoma remain as pelagic larvae between 12 and 24 months (Booth 1994). This prolonged PLD led to the assumption of dispersal over a long geographical area. A larval dispersal model that included larval vertical migration, behavior and larval growth in addition to hydrodynamics determined that there is larval exchange between sub-populations from Victoria, South Australia and Tasmania (Bruce et al. 2007).

Despite studies that have shown that *J. edwardsii* comprises a single stock in Australia (Ovenden et al. 1992; Morgan et al. 2013; Villacorta-Rath et al. 2016), the fishery is managed by each state separately. South Australia, Victoria and Tasmania have similar fisheries management tools in place, however each state develops their management plans independently (Linnane et al. 2010c). The southern fishing zone of South Australia has been hypothesized as a likely source of recruits into the Tasmanian populations due to oceanic features (Bruce et al. 2007). This implies that the Tasmanian rock lobster fishery is influenced by the status of the resource in the South Australian southern zone. The current level of egg production in the southern zone of South Australia is at 9% of its unfished level (Linnane et al. 2014b). This is below the recommended egg production level of 20% of virgin biomass (Australia Department of Agriculture, Fisheries and Forestry 2007) and of great concern for a population that is believed to be sourcing recruits into other areas. While hydrodynamic models generate hypotheses on potential dispersal pathways, molecular tools can test whether these source and sink pathways occur. Hence, using molecular markers it is possible to disentangle the effects of the physical environment on larval dispersal and its consequences on fluctuations in recruitment.

Monitoring of recruitment abundance has been carried out in southeastern Australia since the 1970 using puerulus collectors located in permanent sites (Linnane

et al. 2010c). The abundance of *J. edwardsii* recruits, defined as individuals transferring from the planktonic to the demersal population (Sale 1990), exhibits high year-to-year variability (Booth 1994; Linnane et al. 2014a; Hinojosa et al. 2017). Due to the wide geographical range of this species, environmental factors influence settlement abundance differently from region to region (Hinojosa et al. 2017). Puerulus index, the mean abundance of puerulus per collector per year, has been correlated to amount of commercial catch rate five years later in Tasmania (Gardner et al. 2001). If recruitment variability (and therefore future commercial catch) is dependent on egg production, then it is important to understand the link between these two. This is especially important since the fishery is managed over seven jurisdictions in Australia while constituting a single stock. While there have been studies on settlement abundance and the environmental factors driving it, far less attention has been paid to measuring variability of the genetic diversity in this species. It is also unknown how genetic diversity is coupled to the variability in settlement abundance.

1.4 Applying genomic tools to inform fisheries management

Advances in molecular methods currently allow the identification of hundreds to thousands of molecular markers across the entire genome called single nucleotide polymorphisms (SNPs). SNPs are variations in single nucleotides among individuals (Peterson et al. 2012). The large number of molecular markers allows the possibility of discovering neutral genetic variation between populations where traditional markers failed to detect genetic divergence (Benestan et al. 2015). Moreover, recent migration rates can also be determined with high accuracy using genetic (10-100 molecular markers) and genomic (>1,000 molecular markers) tools (Bernatchez et al. 2017). This is especially important in determining putative source populations in commercial species (Benestan et al. 2015). Finally, scanning the entire genome

enables investigating adaptive loci that exhibit very high genetic differentiation between populations of marine species that lack of divergence at neutral markers (Allendorf et al. 2010). Uncovering local adaptation within populations of commercial species can also help to adjust management practices in order to match demographic and biological management units (Corander et al. 2013).

1.5 Thesis structure

The overall aim of this thesis was to study variability in genetic identity at different spatio-temporal scales. This was done in order to determine the potential biological, environmental (e.g., dispersal history) or adaptive (e.g. natural selection) drivers of population structure. This thesis has been prepared as a series of four manuscripts that constitute four data chapters and that will be submitted for publication in peer-reviewed journals. Using double digest restriction site-associated DNA sequencing (ddRADseq), a next generation sequencing approach, I was able to identify SNP markers for *J. edwardsii* in order to: (1) determine the degree, and potential causes, of population subdivision in *J. edwardsii* at a large geographic scale, (2) investigate the relationship between fluctuations in settlement abundance and genetic diversity of *J. edwardsii*, (3) examine the potential drivers of phenotypic variation in recently settled *J. edwardsii*, and (4) test the predictions of a larval dispersal model of *J. edwardsii* in southeast Australia to determine potential source populations.

Chapter 1. Constitutes a general introduction to the subject of genetic connectivity and the mechanisms behind certain common patterns that cannot be explained by ocean advection only, such as chaotic genetic patchiness.

Chapter 2. This data chapter determines the degree of subdivision at a broad spatial scale (1,000's km), between *J. edwardsii* populations in Australia and New Zealand, and explores local adaptation as a potential cause of genetic structure.

Chapter 3. This data chapter examines the relationship between fluctuations in settlement abundance and genetic diversity of *J. edwardsii* over four settlement years at a medium spatial scale (100's km), between Tasmania and South Australia, and proposes the existence chaotic genetic patchiness in settlers as well as weak post-settlement selection.

Chapter 4. This data chapter investigates dispersal history as well as natural selection as potential drivers of the observed large phenotypic variation in recently settled *J. edwardsii* at a small spatial scale (10's km), on the east coast of Tasmania.

Chapter 5. This data chapter assigns *J. edwardsii* recruits to most likely populations of origin in South Australia and Tasmania.

Chapter 6. Constitutes a general discussion of the findings of this thesis and provides future directions for research on genetic connectivity and demography indicators of *J. edwardsii*.

Chapter 2: Outlier SNPs enable food traceability of the Southern rock lobster, *Jasus edwardsii*

This work has been published as:

Villacorta-Rath C, Ilyushkina I, Strugnell JM, Green BS, Murphy NP, Doyle SR, Hall NE, Robinson AJ, Bell JJ (2016) Outlier SNPs enable food traceability of the southern rock lobster, Jasus edwardsii. Marine Biology, 163:223. doi:10.1007/s00227-016-3000-1

2.1 Abstract

Recent advances in next generation sequencing have enhanced the resolution of population genetic studies of non-model organisms through increased marker generation and sample throughput. Using double digest restriction site-associated DNA sequencing (ddRADseq), we investigated the population structure of the commercially important Southern rock lobster, *Jasus edwardsii,* in Australia and New Zealand with the aim of identifying a panel of SNP markers that could be used to trace country of origin. Four ddRADseq libraries comprising a total of 91 individuals were sequenced on the Illumina MiSeq platform and demultiplexed reads were used to create a reference catalog of loci. Individual reads were then mapped to the reference catalog and variant calling was performed. We have characterized two single nucleotide polymorphism (SNP) panels comprised in total of 656 SNPs. The first panel contained 535 neutral SNPs, and the second, 121 outlier SNPs that were characteristic of being putatively under selection. Both neutral and outlier SNP panels showed significant differentiation between the two countries, with the outlier loci demonstrating much larger F_{ST} values (F_{ST} outlier SNP panel = 0.134, $P < 0.0001$; F_{ST} neutral SNP panel = 0.022, $P < 0.0001$). Assignment tests performed with the outlier

SNP panel allocated 100% of the individuals to country of origin, demonstrating the usefulness of these markers for food traceability of *J. edwardsii*.

2.2 Introduction

Marine benthic invertebrates typically exhibit a pelagic larval phase that serves as a mechanism of dispersal and maintains connectivity between sub-populations (Eckman 1996). It is widely accepted that larvae with a short pelagic larval duration (PLD) are more prone to settling close to their parents, generating genetic structure at broad spatial scales (Palumbi 1994). In contrast, long PLDs can potentially lead to an absence of, or low population structure due to dispersal of larvae over large geographical areas (Shanks et al. 2003). However, most larval transport is largely determined by hydrodynamic features, which can cause strong genetic differentiation, even in species with a relatively long PLD (Palumbi 1994). An increasing number of studies using genetic markers have concluded that larval duration cannot be directly used as a predictor of genetic structure (Shanks 2009; Wei et al. 2013; Teske et al. 2015).

The Southern rock lobster, *Jasus edwardsii*, is distributed from southern Australia and the Tasman Sea to all coasts of New Zealand. *J. edwardsii* is a commercially important species in both countries and fisheries management is carried out independently in Australia and New Zealand. This resource represents a substantial income for economies of both countries, providing annual revenue of approximately US$292 million to Australia (ABARE-BRS 2009) and US$204 million to New Zealand (www.stats.govt.nz). The main export market for both countries is Asia, where lobsters are mostly exported live (ABARE-BRS 2009; Jeffs et al. 2013).

Despite the protracted pelagic larval duration of up to 24 months (Booth and Phillips 1994), modeling simulations of larval trajectories have estimated that only

8% of larvae released from Australia have the potential to reach New Zealand (Bruce et al. 2007). An earlier genetic study that characterized the structure of six rock lobster populations from around New Zealand and two populations in Australia found evidence for restricted gene flow across the species range (Thomas and Bell 2013). The hypothesis of a panmictic *J. edwardsii* population throughout its geographical range was rejected based on significant F_{ST} (F_{ST} = 0.011) from nine microsatellite markers (Thomas and Bell 2013). However, the authors suggested the possibility of larval migration from Tasmania (Australia) to central New Zealand. In support of this hypothesis, a subsequent study conducted with eight microsatellite markers using lobsters collected from Tasmania and the southern zone of the South Island of New Zealand also revealed population structure between countries (Morgan et al. 2013). The presence of genetic structure between Australia and New Zealand populations detected using microsatellites (Thomas and Bell 2013; Morgan et al. 2013) highlights the potential for using genetic markers to assign location of origin to lobsters, which may be useful for fisheries management purposes.

The use of genome-wide SNP markers, in contrast to microsatellite markers, has the potential to improve resolution in the estimation of population structure, migration rates, dispersal and population assignment (Morin et al. 2004; Benestan et al. 2015), as well as the ability to explore genomic regions under selection. Recent studies have identified high levels of population structure when analyzing small numbers of outlier markers in marine fish (Corander et al. 2013; Milano et al. 2014; Candy et al. 2015). For example, 299 neutral SNPs identified large-scale population subdivision of the widespread European hake, *Merluccius merluccius*, between the Atlantic and Mediterranean Seas, but significantly finer scale resolution was found when analyzing just 7 and 19 outlier SNPs within the Atlantic and Mediterranean basins, respectively (Milano et al. 2014). Similarly, fine scale population structure of eulachon,

Thaleichthys pacificus, in North America was distinguished through 193 outlier SNPs, in comparison to lower genetic differentiation detected when analyzing 3911 neutral SNPs (Candy et al. 2015). Both studies attributed the high level of genetic variation in outlier SNPs to local adaptation.

The high levels of population differentiation detected with markers under selection makes them appropriate for traceability of commercial fisheries species (Araneda et al. 2016). Traceability of fish products is being increasingly used for consumer protection and for regulatory enforcement, especially in unreported and unregulated fishing (Ogden 2008). Although the *J. edwardsii* fisheries in Australia and New Zealand are managed sustainably, both countries export lobsters to the Asian market. In the past, China has restricted Australian imports due to public health concerns and economic reasons. Therefore, efficient assignment of commercialized *J. edwardsii* to country of origin could prevent any conflict between Australia and New Zealand if any further bans are imposed in the future.

The aim of this study was to identify a panel of SNP markers that would enable high population assignment success and therefore could be used to trace country of origin for *J. edwardsii* to either New Zealand or Australia. We used a double digest restriction site-associated DNA (ddRADseq) approach (Peterson et al. 2012) to explore genetic structure of *J. edwardsii* using both neutral markers and markers putatively under selection. The high level of genetic differentiation exhibited by the outlier SNP panel allowed us to successfully assign individuals to population of origin.

2.3 Materials and methods

Sample collection

A total of 40 individuals from five sites (corresponding to three regions) in Australia and 48 individuals from four sites (corresponding to four regions) in New Zealand were collected for the present study (Fig. 2.1) between 2011 and 2014 (Table 2.1). Adult lobsters on West Tasmania (AUS), East Tasmania (AUS), the Hauraki Gulf (NZ) and the Chatham Islands (NZ) were caught using commercial baited lobster pots. For the Australian samples, a pleopod clip was taken from each lobster and preserved in 90% ethanol. For the New Zealand samples, lobster legs were removed from live specimens and frozen. Downstream analyses suggest that differences in sample preservation did not produce a batch effect, since samples from the Hauraki Gulf and Chatham islands were assigned into the same cluster as the rest of New Zealand sampling sites (see Results). In the case of Stewart Island (NZ), Tonga Island (NZ) and Merri Marine Sanctuary (AUS), adult lobsters were collected by divers and legs (New Zealand specimens) or pleopod clips (Australian specimens) were taken from each lobster and immediately preserved in 90% ethanol.

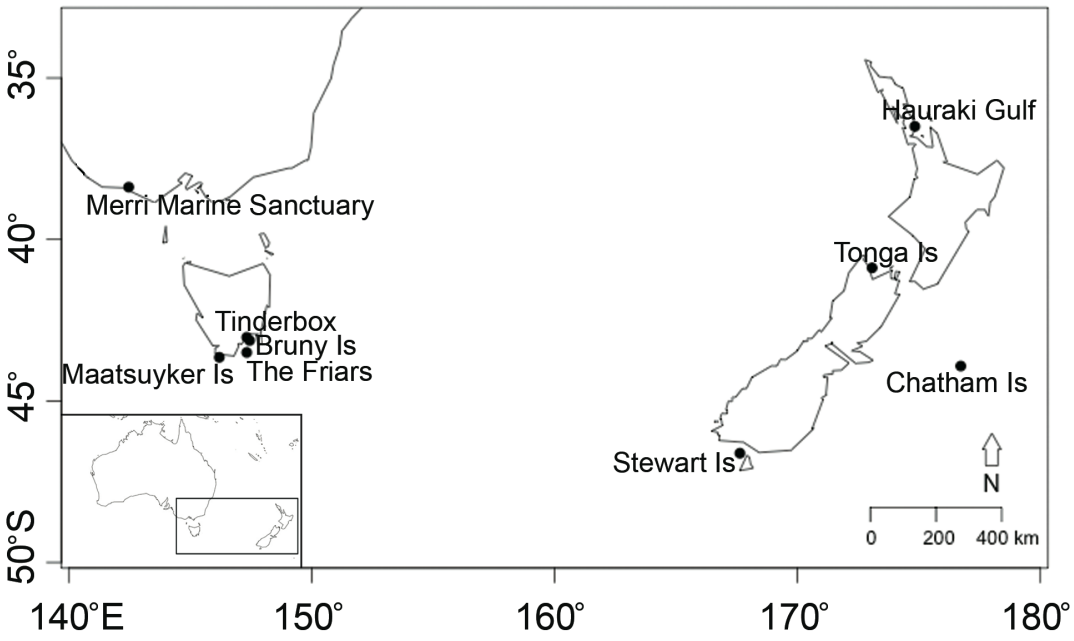


Fig. 2.1 Sampling sites in Australia and New Zealand

Table 2.1 Sampling sites for the *J. edwardsii* collected in Australia and New Zealand. Number of individuals sequenced (n) and final number of individuals (n_f) after filtering for missing data and removal of replicates are reported for each sampling site

Country	Region	Sampling site	Sampling year	Stage	Latitude	Longitude	n	n_f
Australia	Victoria	Merri Marine Sanctuary (MMS)	2013	Adult	38°23'S	142°28'E	8	7
	West Tasmania	Maatsuyker Island (MAA)	2014	Adult	43°39'S	146°12'E	11	9
	East Tasmania	The Friars (FSX)	2014	Adult	43°30'S	147°20'E	8	4
	East Tasmania	Bruny Island (BRU)	2013	Adult	43°08'S	147°27'E	4	3
	East Tasmania	Tinderbox (TXX)	2013	Adult	43°02'S	147°20'E	9	7
New Zealand	North NZ	Hauraki Gulf (HGU)	2011	Adult	36°30'S	174°50'E	13	10
	Central NZ	Tonga Island (TIS)	2013	Adult	40°53'S	173°04'E	11	11
	East NZ	Chatham Islands (CHI)	2013	Adult	43°55'S	176°43'E	12	12
	South NZ	Stewart Island (SIS)	2013	Adult	46°38'S	167°37'E	12	12

DNA extractions

DNA was extracted using the DNeasy Blood and Tissue kit (Qiagen) following the manufacturer's instructions. The DNA concentration of each sample was determined using a Qubit® 2.0 Fluorometer (Life Technologies). DNA integrity was determined through gel electrophoresis and samples with predominantly high molecular weight DNA (corresponding to a band 1,000 base pairs, bp, or higher) were preferentially selected for ddRADseq library preparation.

ddRADseq library preparation and sequencing

A modified version of the ddRADseq protocol developed by Peterson et al. (2012) was used to make multiplexed sequencing libraries (https://molecularbiodiversity.wordpress.com/home/protocols/). Briefly, 250 ng of genomic DNA was digested using the restriction enzyme EcoRI (GAATTC, infrequent cutter) and AciI (CCGC, frequent cutter) at 37°C for 16 h. Subsequently, sequencing adapters containing in-line barcodes were ligated to the sheared DNA fragments, after which low molecular weight DNA fragments as well as non-ligated adapters were removed using a double size selection protocol employing Agencourt AMPure XP magnetic beads (Beckman Coulter) (Lennon et al. 2010). Index sequences (based on TruSeq LT) and flow-cell attachment regions were incorporated by primer extension PCR. PCR products were cleaned using AMPure XP (Beckman Coulter), after which the DNA concentration was standardized, pooled and a gel size selection was performed to obtain a DNA fragments between 400 and 500 bp. DNA was extracted from the gel using the Wizard SV Gel and PCR Clean-Up System (Promega) and the concentration of the final ddRADseq library was determined using Qubit 2.0 fluorometer. All ddRADseq libraries were spiked with 10% PhiX Control

v3 and sequenced on the Illumina MiSeq next generation sequencing platform using v2 2x250 bp kits.

The number of samples to be sequenced in a single run was determined through two pilot sequencing runs. For the first pilot run, 12 individuals from New Zealand were sequenced and for the second pilot run 13 specimens from Australia were sequenced. We determined the number of polymorphic loci and the number of reads and sequencing depth per individual and concluded that ddRADseq libraries consisting of 45 individuals would yield sufficient coverage and depth of loci to be sequenced in a single run. Subsequently, two more ddRADseq libraries were sequenced, for a total of four ddRADseq datasets.

Technical replicates

Technical replicates were included in the second, third and fourth sequencing libraries prepared. This was particularly important given that our ddRADseq libraries were not prepared and sequenced at the same time, which may have had introduced technical- and sequencing-derived differences between libraries (Mastretta-Yanes et al. 2015). A principal component analysis (PCA) was used to visualize the spatial distribution of replicates and identify whether there was a batch effect due to library preparation and sequencing that could bias results. The PCA was performed using the R package adegenet v.1.4-1 (Jombart and Ahmed 2011). A summary of the distribution of technical replicates among libraries is given in Table S1 (Supporting information).

Preliminary analyses of raw sequencing data

Quality of the reads was initially examined using the FastQC v.0.10.1 quality control tool (Babraham Bioinformatics). Subsequently, uniquely indexed and

barcoded samples were demultiplexed using the "process_radtags" protocol from Stacks v.1.29 (Catchen et al. 2011). Based on the FastQC report, sequences were trimmed to 75 bp to assure that the Phred Quality Score (Q) of all reads were above 30. Trimmed reads were assessed for bacterial and viral contamination using Kraken v.3.5.0 (Wood and Salzberg 2014). This software compares sequence reads against a database to identify reads that match the taxonomic groups present in the database. Reads that do not match those of the database are output as "unclassified", or non-bacterial or viral reads, which were extracted from the raw trimmed data for further analysis. In the absence of a reference genome, removal of bacterial or viral reads is important since it is impossible to determine whether a sequence belongs to the study organism or to a contaminant, providing biased results in downstream analyses (Merchant et al. 2014).

Reference catalog building, alignment and variant calling

Since *J. edwardsii* is a non-model species, we identified a catalog of the most frequently sequenced loci using the "rad-loci" pipeline (https://github.com/molecularbiodiversity/rad-loci) developed at La Trobe University. Sequence data from all individuals was first pooled and then clustered using VSearch v.1.1.3 (Rognes et al. 2016) to identify putative loci. Initially only clusters of reads with a depth of at least 103 (and therefore likely to appear at least once in most of the samples) were retained. This means that the number of raw reads required to form a cluster was 103. Sequences were considered to be sufficiently related if they shared at least 94% identity (4 bp maximum difference in a 75 bp read, allowing single nucleotide variations and indels). Therefore, the allowed number of mismatches between two clusters was set to 5bp. Assuming that each member of the cluster was an allele, only clusters that were composed by a minimum of two members and a

maximum of 16 members were kept. The minimum number was based on the fact that we wanted to obtain bi-allelic data. A second round of clustering of the remaining reads at 94% identity was performed, followed by another filtering of clusters that were not comprised by a minimum of two and a maximum of 16 members. After re-filtering, individual samples were mapped back to the filtered clusters, each cluster was now called a "locus" and it consisted of one representative sequence and up to 16 variations. Finally, samples were mapped back to the identified loci allowing for a maximum of 40% missing data across samples in each locus. The total number of "reference" loci identified by this pipeline was 1,054 (Table S2, Supporting information).

Individual reads were subsequently mapped to the reference loci catalog using the software Bowtie2 v.0.7.12 (Langmead and Salzberg 2012). Finally, variant calling of mapped loci was performed using the Genome Analysis Toolkit (GATK) v.3.3_0 (McKenna et al. 2010), yielding a total of 947 SNPs (Table S2, Supporting information).

SNP filtering

The putative RAD loci were filtered to ensure that known confounding variables, such as non bi-allelic loci, missing loci, allele dropout (ADO) and linkage disequilibrium (LD) (Henning et al. 2014), were minimized prior to population level analyses using VCFtools v.0.1.13 (Danecek et al. 2011). Only bi-allelic loci were retained, using the options --min-alleles 2 --max-alleles 2. Additionally we selected a single SNP per locus with the option --thin 75, given that loci were 75 bp long. Since paired-end sequencing was performed and therefore pairs of loci of 75 bp length could potentially be in LD, we set a pairwise LD measure threshold of $r^2 > 0.8$ to remove potentially linked loci using the option --min-r2. Average coverage was 44,

allowing for a minimum coverage of 5x to minimize ADO. Rare alleles were also removed by setting a minor allele frequency of 0.1 with the option --maf. Finally, both loci and individuals with more than 20% missing data were excluded from the analysis, yielding a final total number of 656 SNPs (Table S2, Supporting information).

SNP characterization

Detection of neutral loci and loci putatively under selection was performed using LOSITAN (Beaumont and Nichols 1996; Antao et al. 2008) using 100,000 simulations, a confidence interval of 0.99, and a false discovery rate of 0.1 (Jacobsen et al. 2014). LOSITAN uses an F_{ST}-outlier approach that identifies loci as outliers when their F_{ST} values are too high or too low compared to neutral expectations (Antao et al. 2008). Loci found to be under balancing and positive selection will be hereafter referred as outlier loci. All downstream analyses were performed and are reported for each SNP panel separately.

Finally, to examine whether sequence reads aligned to protein coding regions, demultiplexed untrimmed reads were screened through tBLASTx v.2.2.29+ (Altschul et al. 1997). This program searches a translated nucleotide database with putative translated nucleotide queries. Raw reads (150 bp) were used rather than the 75 bp reference loci to improve the BLAST alignment length and hence specificity. Subsequently, queries with statistically significant e-values ($E < 0.01$, Karlin and Altschul, 1990) were screened against the reference loci using BLASTn v.2.2.29+ (Altschul et al. 1997) to identify any loci that were contained on those reads and therefore linked to those genes.

Analyses of genetic diversity

The level of observed (H_o) and expected heterozygosity (H_E) in each population, as well as F statistics for both SNP panels, were calculated using the R packages adegenet v.1.4-1 (Jombart and Ahmed 2011) and pegas v.0.8-2 (Paradis 2010). A two-sample Wilcoxon test was used to detect whether mean H_o differed significantly from mean H_E for both SNP panels.

Effective population size estimation

Effective population size (Ne) was estimated by the software NeEstimator v.2.01 (Do et al. 2014) using the linkage disequilibrium model based on allele frequencies of all 656 loci. This model provides the most precise estimation of Ne among other single-sample methods (Waples and Do 2010; Do et al. 2014).

Analysis of population structure

Population structure was investigated through discriminant analysis of principal components (DAPC). DAPC was performed and results were plotted using the R package adegenet v.1.4-1 (Jombart and Ahmed 2011). This analysis assigns individuals to clusters and selects the best number of clusters based on Bayesian Information Criterion (BIC) (Jombart et al. 2010). The results are comparable to those obtained by STRUCTURE (Pritchard et al. 2000), with the advantage that DAPC explores genetic structure without making assumptions about the genetic model of the study population (Jombart et al. 2010).

Assignment of individuals to country of origin

Assignment tests were performed using the program Geneclass2.0 (Piry 2004) to test the effectiveness of the outlier SNP panel to assign individuals to country of

origin. We simulated $10x10^3$ multilocus genotypes using the algorithm described by Paetkau et al. (2004). We used the Bayesian allele frequency estimation method of Rannala and Mountain (1997) whereby each individual is removed from the baseline and assigned to the most likely population.

2.4 Results

Variant calling and SNP filtering

ddRADseq libraries were prepared from seven sub-populations of the Southern rock lobster from Australia and New Zealand and sequenced using an Illumina MiSeq. A "reference" catalog of loci was developed from a total of 88 individuals using an average of 580,000 reads per individual. After variant calling and SNP filtering, our final dataset comprised 75 individuals, including 30 Australian and 45 New Zealand samples. From this data, a total of 954 SNPs were identified, and after filtering (MAF = 0.1, LD < 0.8 between loci, 1 SNP per loci, < 20% missing data of loci and individuals), a total of 656 SNPs were obtained. Using the software LOSITAN to discriminate between putatively neutral variants and variants characteristic of being under selection, two panels of SNPs were identified consisting of 535 neutral and 121 outlier SNPs respectively.

Consistency in assignment of technical replicates

Technical replicates were included in three ddRADseq libraries. The PCA showed consistency in assigning technical replicates from the same individual close to each other and within the cluster of each country. This means there was no batch effect produced by different library preparation and sequencing runs (Fig. S1, Supporting information).

Genetic diversity and effective population size

Mean expected and observed heterozygosity for both countries were higher for the neutral SNP panel than for the outlier panel and the New Zealand population exhibited higher heterozygosity than the Australian population (Table 2.2). The negative values of the inbreeding coefficient (F_{IS}) estimated for both SNP panels and both countries are indicative of heterozygote excess. The two-sample Wilcoxon test indicated that H_O was significantly higher than H_E in both SNP panels for the New Zealand population (P < 2.2e-16 and P = 0.0199, respectively for the neutral and outlier SNP panels). For the Australian population, significant differences between H_O and H_E were only found for the neutral SNP panel (P < 2.2e-16 and P = 0.07222, respectively for the neutral and outlier SNP panels).

Estimation of effective population size in Australia and New Zealand gave infinite values, with confidence intervals of 1334.6 - ∞ for Australia and infinite for the New Zealand *J. edwardsii* population. Infinite values of estimated Ne and confidence intervals may be due to larger than expected sampling error (Do et al. 2014).

Table 2.2 Descriptive statistics for Australia and New Zealand given by the neutral and outlier SNP panels

	N	N_{loci}	H_O (± SD)	H_E (± SD)	F_{IS}	F_{ST}	*P*
Neutral SNP panel							
New Zealand	45	535	0.633 (± 0.279)	0.399 (± 0.118)	-0.588		
Australia	30	535	0.518 (± 0.268)	0.369 (± 0.126)	-0.401	0.022	<0.0001
Outlier SNP panel							
New Zealand	45	121	0.386 (± 0.319)	0.264 (± 0.176)	-0.463		
Australia	30	121	0.355 (± 0.271)	0.273 (± 0.169)	-0.297	0.134	<0.0001

N sample size, N_{loci} number of SNPs, H_O mean observed level of heterozygosity, H_E mean expected level of heterozygosity, F_{IS} fixation index (inbreeding coefficient), F_{ST} pairwise fixation index between Australia and New Zealand

Population structure

While both panels detected population structure between Australia and New Zealand, the outlier loci detected greater differentiation between countries (Table 2.2). DAPC analysis using both SNP panels detected two clusters, each of them representing each of the two countries (Fig. 2.2). However, the higher genetic divergence given by the outlier SNP panel was detectable in the distance between both clusters, which was one order of magnitude higher than that of the neutral SNP panel (Fig. 2.2c, a, respectively). To determine the outlier SNPs causing most of the differentiation between countries we inspected the associated allele loadings (Jombart et al. 2010), showing that 30 SNPs contribute to the discrimination between Australia and New Zealand (Fig. S2, Supporting information).

Assignment of individuals to country of origin

The outlier SNP panel correctly assigned 100% of individuals to population of origin, with a quality index of 87.72% (Table S3, Supporting information).

Characterization of catalog loci

An analysis of the ddRAD loci to identify coding sequence variants using tBLASTx did not produce any significant alignment with loci containing neutral or outlier SNPs.

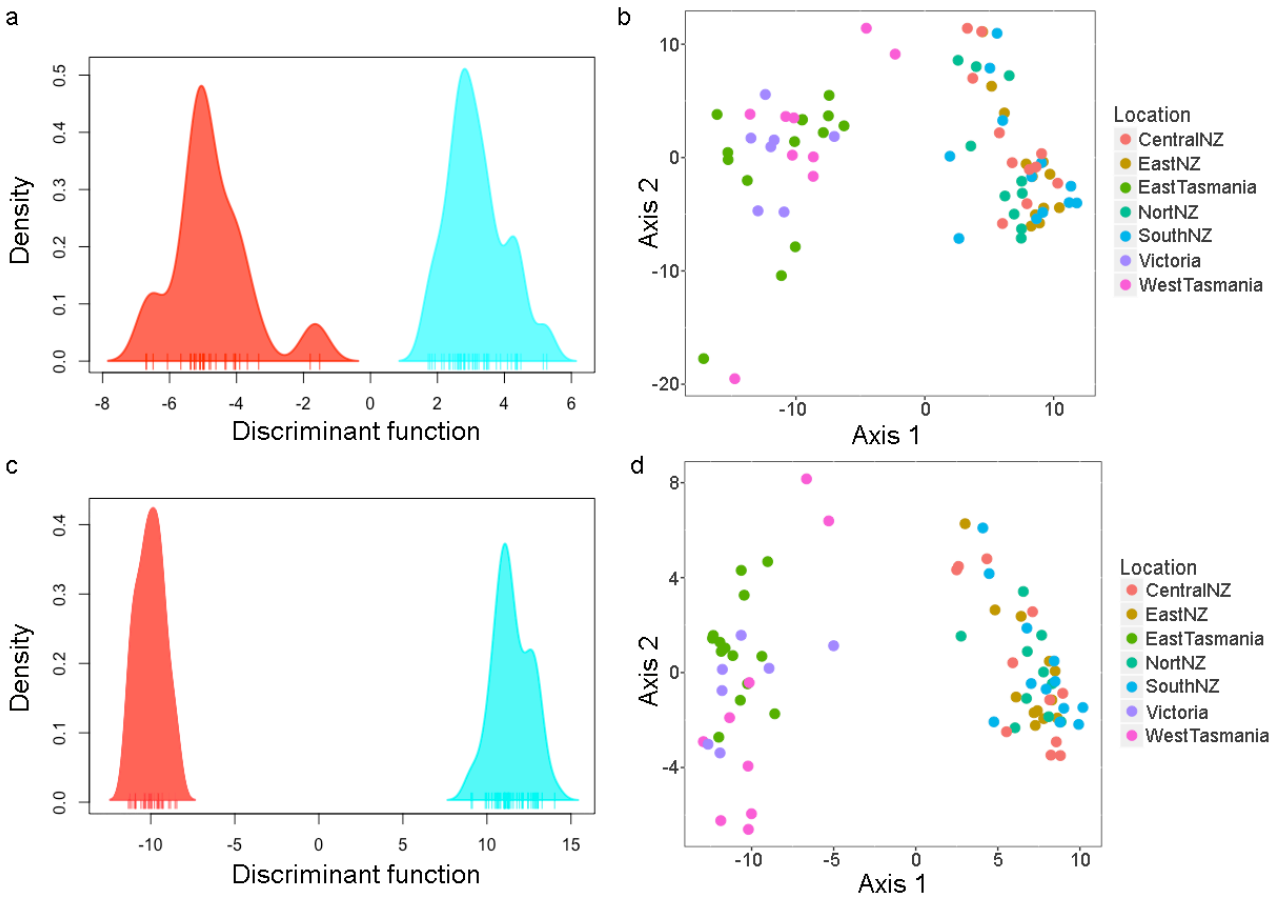


Fig. 2.2 First principal component resulting from the discriminant analysis of principal components (DAPC) using (a) the neutral SNP panel and (c) the outlier SNP panel. The red group represents 30 Australian individuals while the blue group represents 45 New Zealand individuals. Principal component 1 and 2 resulting from a Principal Component Analysis (PCA) using (b) the neutral SNP panel and (d) the outlier SNP panel. Sampling regions are represented by colors

2.5 Discussion

This study represents the first genome-wide population genetic analysis of a rock lobster species. We produced a panel of outlier and neutral loci to investigate population structure between Australia and New Zealand and the potential for assignment of individuals to population of origin. The Southern rock lobster, *J. edwardsii*, is a commercially important species in Australia and New Zealand, and, given reported evidence of differing degrees of genetic connectivity (Thomas and Bell 2013; Morgan et al. 2013), further fine scale genetic mapping of populations from both countries is required to understand the extent of connectivity between countries.

As seen in other marine species, our outlier SNP panel showed greater genetic differentiation between Australia and New Zealand than the neutral panel. Therefore we propose that the outlier SNP panel has potential as an effective method for determining population of origin in *J. edwardsii*.

Evidence for a population bottleneck

Significant differences between mean H_O and H_E were detected for the neutral SNP panel in both countries and the outlier SNP panel in New Zealand, and the negative values of the fixation index indicates heterozygote excess. Heterozygote excess can be a result of low depth of coverage, since in the presence of low depth data homozygote genotypes could be erroneously called heterozygous (Nielsen et al. 2012b). However, we do not consider this to be the case in the present study since our average depth of coverage is 43 and it has been suggested that a minimum of 5x coverage gives reliable estimates of allele frequencies (Ellegren 2014).

Heterozygote excess may also be the result of a small effective population size caused by a past population bottleneck many generations ago (Luikart and Cornuet 1998) or as a result of a relatively small number of individuals contributing to each generation (Hedgecock and Pudovkin 2011). However, heterozygote excess caused by a population bottleneck can be a transient state (Luikart and Cornuet 1998). Thomas and Bell (2013) estimated an effective population size of 2,717 in Australia and 1,189 in New Zealand. These estimates are low and therefore it is possible that the populations have undergone a bottleneck. The infinite estimates of Ne in the present study do not provide any evidence of genetic drift due to small Ne (Do et al. 2014). However, as discussed above, our infinite values of estimated Ne and confidence intervals may be due to larger than expected sampling error. When Ne is large or if

there is limited data available, the estimate of Ne will be negative and the biological interpretation is that Ne = ∞ (see Waples and Do, 2010).

Performance of the neutral SNP panel for detecting population structure

Studies on population structure of marine organisms based on neutral markers typically demonstrate significant but weak genetic differentiation (Nielsen et al. 2009b; Milano et al. 2014). Neutral markers are affected by demography and the evolutionary history of populations; therefore they evolve as a result of genetic drift and migration (Luikart et al. 2003). For this reason, even low levels of migrant exchange can maintain genetic homogeneity between populations over long periods of time (Cano et al. 2008; Allendorf et al. 2010). The neutral SNP panel used in the present study provided evidence for significant but weak genetic differentiation between countries, in accordance with the findings of microsatellite markers, which displayed similar levels of genetic differentiation between Tasmania and New Zealand (F_{ST} = 0.029) (Morgan et al. 2013).

Performance of the outlier SNP panel for detecting population structure and implications for fisheries management

In the present study, the much higher and statistically significant F_{ST} value exhibited by the outlier SNP panel (F_{ST} = 0.134) demonstrated that loci putatively under selection have high power for detecting genetic structure between Australia and New Zealand and therefore could be used for determining country of origin in *J. edwardsii*.

Molecular techniques for seafood authentication are increasingly used to monitor fish stocks that are still commercially viable but are becoming threatened by overfishing, and to protect the consumer from fraudulent practices (Ogden 2008;

Sorenson et al. 2013; Larraín et al. 2014). DNA barcoding is one of the preferred techniques due to its applicability to degraded material, low DNA requirement, simple protocol, time efficiency and reproducibility (Wong and Hanner 2008). However, DNA barcoding is more effective for inter-specific differentiation, since it targets the mitochondrial cytochrome *c* oxidase I (COI) gene which can be highly conserved between subpopulations of the same species (Ogden 2008). In contrast, techniques that target the nuclear genome provide the potential for intra-species assignment to population of origin (Nielsen et al. 2012a).

Trials using microsatellite markers for food traceability have been carried out with limited success in other commercial marine species. For example, a panel of nine microsatellites produced up to 50% of correct assignments to country of origin of the mussel *Mytilus chilensis* in the presence of a global F_{ST} lower than 0.042 (Larraín et al. 2014). These authors suggested that having more informative loci and using SNP markers could improve assignment success. In contrast, higher success (92%) was obtained when combining 13 microsatellite markers with data from the mitochondrial control region to accurately assign individuals to ocean of origin in Atlantic and Pacific blue marlin stocks with low genetic differentiation ($F_{ST} < 0.01$) (Sorenson et al. 2013). However, the suitability of loci putatively under selection for more reliable assignment to population of origin has been recognized (Martinsohn et al. 2009; Nielsen et al. 2012a) and recent studies have successfully demonstrated the power of outlier SNPs for this purpose (Araneda et al. 2016).

The outlier SNP panel obtained in the present study demonstrated strong genetic differentiation of *J. edwardsii* between Australia and New Zealand ($F_{ST} = 0.134$). Based on these markers, complimentary high throughput technologies that rely on prior sequence information, such as target capture or loop-mediated isothermal amplification (LAMP assay) (Tomita et al. 2008) could be used to differentiate

lobsters caught in New Zealand from those caught in Australian waters in order to avoid mislabeling of country of origin. In particular, LAMP assay amplifies specific regions of the DNA with high specificity, efficiency, rapidity and low cost for preparation and visualization of results (Tomita et al. 2008). This assay is being increasingly used for clinical diagnosis of infectious diseases in developing countries since it does not require expensive laboratory equipment (i.e. a thermal cycler) and it can be performed in 1 hour (Parida et al. 2008). For the particular case of *J. edwardsii*, primers could be specifically designed to target regions of the genome containing the outlier SNPs with highest loadings identified in Fig. S2 (Supporting information) that account for most of the divergence between Australia and New Zealand (Martinsohn et al. 2009).

Genetic structure detected using outlier loci could suggest differences at regions of the genome putatively subject to selection. This divergence could be due to local adaptation to environmental conditions or to post-settlement mortality of unsuited genotypes (Holt and Gaines 1992; Caley et al. 1996; Marshall et al. 2010). Local adaptation to environmental conditions will result in genetic divergence between populations in the presence of high self-recruitment (Holt and Gaines 1992; Sanford and Kelly 2011). Herein, we demonstrated very high levels of self-recruitment within each country using outlier loci. Also, post-settlement mortality of unsuited genotypes will preserve the local genetic pool by removing migrants that are not fit to survive under particular environmental conditions (DeWitt et al. 1998; Marshall et al. 2010). Currently it is impossible to unravel the degree at which self-recruitment and post-settlement mortality are contributing to genetic differentiation of *J. edwardsii* between Australia and New Zealand and this is beyond the scope of this study.

Based on the high genetic divergence resulting from the outlier SNP panel, we hypothesize that local conditions may have helped shape patterns of genetic diversity

within adaptive regions of the genome in these populations, as shown in other studies (see Corander et al. 2013; Fraser et al. 2014), however in the absence of any reference genes or transcriptome for *J. edwardsii* or closely related species it is impossible to ascertain this. We can only speculate that differences in environmental conditions between sampling sites may be driving differences at the outlier loci. Empirical evidence suggests that *J. edwardsii* are adapted to specific local conditions, but there is also extensive evidence of very high phenotypic plasticity. For example, growth rates in this species are highly variable and are mainly determined by temperature, density and food availability (Annala and Bycroft 1985; Jeffs and James 2001). Site-specific differences in carapace coloration and growth rates of this species have been reported in Australia (Punt et al. 1997; McGarvey et al. 1999). Individuals inhabiting deep waters are white-colored due to a diet with low concentration of carotenoid pigments and lower nutritional value, which can also impact growth negatively (McGarvey et al. 1999). Translocation experiments of white-colored lobsters into shallow areas demonstrated a change in coloration, growth rates and body condition after 12 months of translocation (Chandrapavan et al. 2010, 2011; Green et al. 2010). Therefore, even when self-recruitment could help retain locally adapted genotypes, phenotypic plasticity can also act to promote growth of *J. edwardsii* in certain environments.

Greater density of SNPs, together with improved but as yet unavailable genetic resources for *J. edwardsii* or closely related species (such as reference genome and transcriptome datasets), would provide further insight into potential genetic evidence of adaption to local environments. Seascape genetics could also help coupling local environmental conditions to genetic distance in order to explain patterns of genetic divergence between populations (Saenz-Agudelo et al. 2015; Giles et al. 2015).

2.6 Conclusions

In this study, we used ddRADseq and Illumina MiSeq next generation sequencing to explore the genetic connectivity of Southern rock lobster populations from Australia and New Zealand. Data from two SNP panels are presented, from which a panel of 121 outlier markers have allowed us to identify clear genetic structure in *J. edwardsii* populations between the two countries as well as high levels of self-recruitment. In addition, highly significant F_{ST} values estimated from the outlier SNP panel could be indicative of local adaptation driving the genetic differentiation between countries. This is particularly important in a commercial species managed by different agencies. Therefore, the outlier SNP panel developed in the present study could be used to differentiate New Zealand from Australian lobsters and therefore be useful for food traceability. We believe that more extensive sampling, including sites along the whole distribution of the species, could identify source and sink regions within each country precisely, which would also help management decision-making in Australia and New Zealand. Finally, continued development of genomics resources, such as transcriptome sequencing, gene characterization and quantitative trait locus discovery is needed in order to explore the link between genotype, phenotype and the environment.

2.7 Acknowledgements

We would like to thank Gary Carlos (University of Tasmania), Colin Fry (University of Tasmania), Daniel Ierodiaconou (Deakin University), Andrew Kent and Kent Way for field assistance and sample collection in Australia. Thanks to Daryl Sykes (New Zealand Rock Lobster Industry Council) for organizing all sample collections in New Zealand and Don Nelson (New Zealand Rock Lobster Industry Council) and Dr. Debbie Freeman (Department of Conservation, New Zealand) for

collecting samples in New Zealand. Thanks for lab assistance to Mel Best, Adam Smolenski and Cecilia Carrea (University of Tasmania). We also thank Michael Amor and Laura Woodings (La Trobe University) who helped developing the ddRADseq protocol and the rad-loci pipeline, respectively. Special thanks to Karen J Miller who contributed to the original project idea. We would like to thank the editor and two reviewers for their constructive suggestions.

2.8 Supplementary information

Table S1 Number of inter- and intralibrary technical replicates

Library #	Intralibrary technical replicates	Interlibrary technical replicates
1	NA	NA
2	NA	1
3	15	5
4	NA	4

Table S2 Number of sequences retained at each step of the rad-loci pipeline and subsequent variant calling and SNP after final filtering for minor allele frequency, linkage disequilibrium, missing data and keeping one SNP per locus

Sequences with ≥103 copies	43,959
Number of clusters	27,260
Number of clusters with 2-16 members	3,102
Total number of alleles after filtering	9,084
Number of loci	2,705
Number of filtered loci	1,054
Number of SNPs after variant calling	947
Number of SNPs after final filtering	656

Table S3 Probability of individuals of belonging to population of origin obtained from Geneclass2

Assigned Sample	Country of origin	Probability of belonging to Australia	Probability of belonging to New Zealand
BRU_01_01_01_I02_R02	Australia	0.9963	0.0001
BRU_03_01_01_I02_R08	Australia	0.9934	0
BRU_05_01_01_I02_R18	Australia	0.9994	0
FSX_06_01_01_I07_R09	Australia	1	0
FSX_07_02_04_I02_R09	Australia	1	0.0001
FSX_09_01_01_I08_R44	Australia	0.9994	0
FSX_12_01_01_I07_R46	Australia	0.9999	0
MAA_07_01_01_I04_R09	Australia	0.9996	0
MAA_09_01_01_I10_R09	Australia	0.9685	0
MAA_11_01_01_I02_R24	Australia	1	0
MAA_13_01_01_I02_R44	Australia	1	0.0002
MAA_14_01_01_I02_R46	Australia	0.9892	0.0001
MAA_15_01_01_I04_R48	Australia	1	0
MAA_54_01_01_I02_R19	Australia	0.9987	0.0035
MAA_67_01_01_I02_R20	Australia	0.9852	0.02
MAA_81_01_01_I02_R29	Australia	1	0
MMS_12_01_01_I09_R09	Australia	1	0
MMS_13_01_01_I07_R14	Australia	0.981	0.0015
MMS_14_02_03_I10_R44	Australia	1	0
MMS_15_01_01_I07_R24	Australia	1	0.0001
MMS_17_01_01_I07_R44	Australia	1	0
MMS_21_01_01_I08_R09	Australia	0.9999	0
MMS_22_01_01_I08_R14	Australia	0.9484	0
TXX_03_01_05_I04_R24	Australia	1	0
TXX_19_01_02_I10_R24	Australia	1	0.0009
TXX_30_01_01_I10_R46	Australia	1	0
TXX_31_01_01_I01_R48	Australia	1	0
TXX_35_01_01_I01_R09	Australia	0.995	0
TXX_36_01_01_I10_R14	Australia	1	0
TXX_48_01_01_I01_R18	Australia	0.993	0
CHI_01_01_01_I04_R05	New Zealand	0.2661	0.9778
CHI_02_01_01_I04_R06	New Zealand	0.3522	0.9993
CHI_03_01_01_I04_R08	New Zealand	0.3776	0.8858

Table S3 continued

Assigned Sample	Country of origin	Probability of belonging to Australia	Probability of belonging to New Zealand
CHI_04_01_01_I02_R02	New Zealand	0.1758	0.9993
CHI_05_01_01_I02_R03	New Zealand	0.2364	0.8927
CHI_06_01_01_I02_R04	New Zealand	0.2369	0.9036
CHI_08_01_01_I02_R11	New Zealand	0.1764	0.9555
CHI_09_01_01_I02_R12	New Zealand	0.2011	0.988
CHI_10_01_01_I02_R18	New Zealand	0.257	0.9999
CHI_11_01_01_I02_R21	New Zealand	0.2354	1
CHI_12_01_01_I02_R19	New Zealand	0.2167	0.9941
CHI_13_01_01_I02_R20	New Zealand	0.1304	0.887
HGU_05_01_01_I04_R18	New Zealand	0.1369	0.9979
HGU_06_02_02_I04_R19	New Zealand	0.2926	0.984
HGU_37_01_01_I04_R20	New Zealand	0.2334	0.9998
HGU_38_01_01_I04_R21	New Zealand	0.2175	0.9998
HGU_39_01_01_I01_R08	New Zealand	0.3165	1
HGU_41_01_01_I04_R02	New Zealand	0.2746	0.9993
HGU_42_01_01_I04_R03	New Zealand	0.2431	0.9238
HGU_43_01_01_I04_R04	New Zealand	0.316	0.3722
HGU_49_01_01_I03_R11	New Zealand	0.2572	0.9035
HGU_50_01_01_I04_R12	New Zealand	0.1681	0.9472
SIS_03_01_01_I03_R18	New Zealand	0.2402	1
SIS_04_01_01_I01_R19	New Zealand	0.2124	0.9997
SIS_20_01_01_I01_R20	New Zealand	0.2514	0.999
SIS_21_01_01_I01_R21	New Zealand	0.2419	0.5949
SIS_47_01_01_I04_R11	New Zealand	0.2754	1
SIS_48_01_01_I03_R02	New Zealand	0.2924	0.7345
SIS_50_01_01_I03_R03	New Zealand	0.3179	0.9366
SIS_51_01_01_I03_R04	New Zealand	0.1125	0.9843
SIS_52_01_01_I03_R05	New Zealand	0.1457	0.9984
SIS_54_01_01_I03_R08	New Zealand	0.101	0.812
SIS_55_01_01_I03_R09	New Zealand	0.0898	0.9989
SIS_56_01_01_I03_R12	New Zealand	0.1022	0.9989
TIS_01_01_01_I01_R02	New Zealand	0.2285	0.9999
TIS_02_01_01_I01_R03	New Zealand	0.2778	0.9568
TIS_03_02_02_I01_R04	New Zealand	0.2393	0.9972

Table S3 continued

Assigned Sample	Country of origin	Probability of belonging to Australia	Probability of belonging to New Zealand
TIS_04_01_01_I01_R05	New Zealand	0.1573	0.9999
TIS_05_01_01_I01_R06	New Zealand	0.2343	1
TIS_08_01_01_I01_R11	New Zealand	0.1834	1
TIS_09_01_01_I01_R12	New Zealand	0.1459	0.9996
TIS_10_01_01_I01_R18	New Zealand	0.2296	1
TIS_12_01_01_I03_R19	New Zealand	0.3009	0.7328
TIS_15_01_01_I03_R21	New Zealand	0.3455	0.4463
TIS_16_01_01_I03_R20	New Zealand	0.4505	0.8366

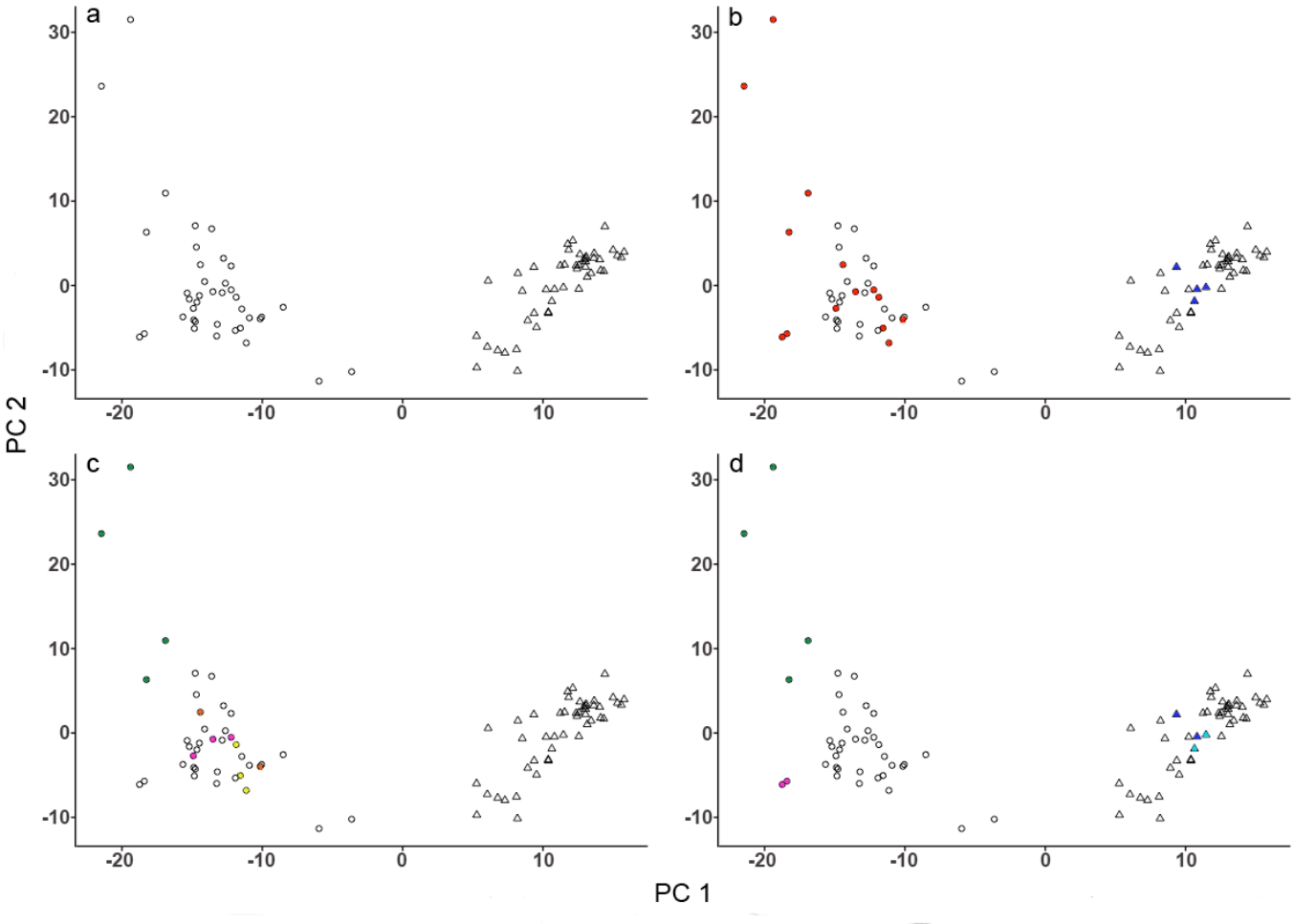


Fig. S1 Principal component analysis based on allele frequencies of all samples from Australia and New Zealand. (a) Circles represent Australian population and triangles represent New Zealand population. (b) Filled figures represent technical replicates within the Australian and New Zealand populations. (c) Filled figures represent technical replicates sequenced in the same MiSeq (intralibrary technical replicates). Each color represents technical replicates from a different sampling site. (d) Filled figures represent technical replicates sequenced in different MiSeq (interlibrary technical replicates). Each color represents technical replicates form a different sampling site

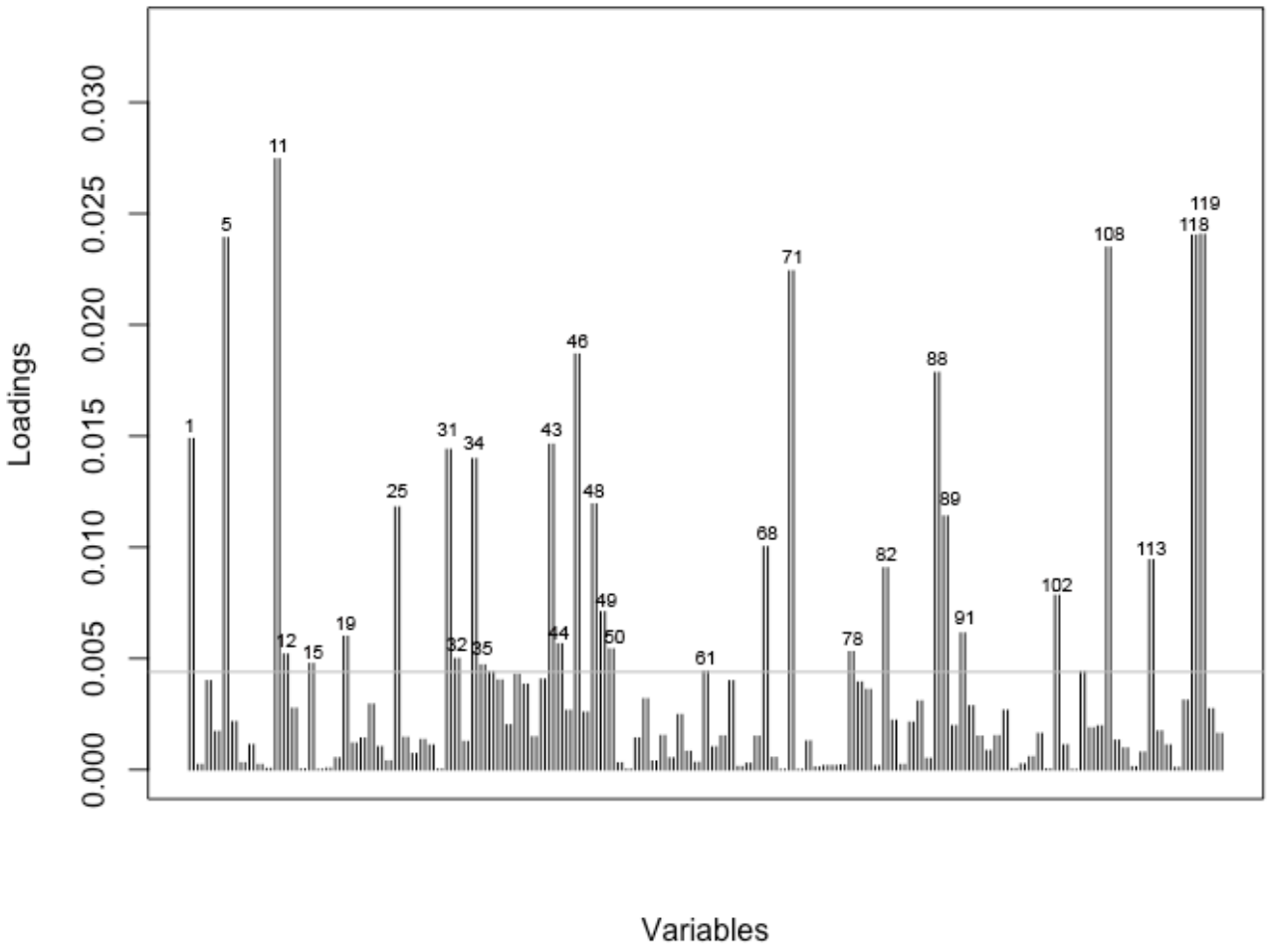


Fig. S2 Contribution of outlier SNPs to the first principal component of the DAPC resulting from the outlier SNP panel. The height of each bar is proportional to the contribution and the dotted line is the threshold above which variables are identified. Only variables above the third quartile threshold are labeled

Supplemental Materials and methods

Migration rates

Recent migration rates between Australia and New Zealand were inferred with a subset of 86 neutral SNPs that exhibited the highest F_{ST} values ($F_{ST} > 0.05$) in the program BayesAss v.3.0.4 (Wilson and Rannala 2003) using a Markov chain Monte Carlo (MCMC) for 10 x 10^6 iterations with a burn-in length of 10^6 iterations.

Supplemental Results

Migration rates

The subset of 86 neutral SNPs with F_{ST} values larger than 0.05 revealed 98% self–recruitment in Australia and 67% in New Zealand. According to these results, approximately 33% of the New Zealand populations are migrants from Australia as well as approximately 0.02% of individuals from the Australian population originated in New Zealand (Table S4, Supporting information).

Table S4 Migration rates between Australia and New Zealand resulting from 86 neutral SNPs (±SD) with $F_{ST} > 0.05$. Rows are source populations and columns are sink populations

	Australia	New Zealand
Australia	0.9792(±0.0140)	0.0208(±0.0140)
New Zealand	0.3262(±0.0070)	0.6738(±0.0070)

Chapter 3: Temporal genetic patterns of diversity and structure evidence chaotic genetic patchiness in a spiny lobster

---

This work has been published as:

Cecilia Villacorta-Rath, Carla A. Souza, Nicholas P. Murphy, Bridget S. Green, Caleb Gardner, Jan M. Strugnell (2018) Molecular Ecology 27, 54-65. doi: 10.1111/mec.14427

3.1 Abstract

Population structure of many marine organisms is spatially patchy and varies within and between years, a phenomenon defined as chaotic genetic patchiness. This results from the combination of planktonic larval dispersal and environmental stochasticity. Additionally, in species with bi-partite life, post-settlement selection can magnify these genetic differences. The high fecundity (up to 500,000 eggs annually) and protracted larval duration (12-24 months) and dispersal of the Southern rock lobster, *Jasus edwardsii*, make it a good test species for chaotic genetic patchiness and selection during early benthic life. Here we used double digest restriction-site associated DNA sequencing (ddRADseq) to investigate chaotic genetic patchiness and post-settlement selection in this species. We assessed differences in genetic structure and diversity of recently settled pueruli across four settlement years and between two sites in southeast Australia separated by approximately 1,000 km. Post-settlement selection was investigated by identifying loci under putative positive selection between recently settled pueruli and post-pueruli and quantifying differences in the magnitude and strength of the selection at each year and site. Genetic differences within and among sites through time in neutral SNP markers indicated chaotic genetic patchiness. Recently settled pueruli at the southernmost site exhibited

lower genetic diversity during years of low pueruli catches, further supporting this hypothesis. Finally, analyses of outlier SNPs detected fluctuations in the magnitude and strength of the markers putatively under positive selection over space and time. One locus under putative positive selection was consistent at both locations during the same years, suggesting the existence of weak post-settlement selection.

3.2 Introduction

Population genetic structure of benthic marine invertebrates is often shaped by larval dispersal and natural mortality during early life history (Li and Hedgecock 1998). The level of genetic connectivity achieved through larval dispersal is critical to our understanding of population dynamics and management of marine species (Cowen et al. 2000). However, dispersal does not always happen in a consistent pattern, as it depends on a suite of factors including ocean advection (Cowen et al. 2000), time of spawning (Hendry et al. 1999), individual reproductive success (Hedgecock 1994), pelagic larval duration (PLD, Palumbi 1994), larval behavior (Raimondi and Keough 1990) and natural selection (Cowen et al. 2000). Variability in the interaction of all those factors means that ephemeral fine-scale genetic heterogeneity or chaotic genetic patchiness could occur rather than the homogenous genetic diversity that would be expected in species with high reproductive output and long PLDs (Johnson and Black 1982).

Selective processes taking place before and after settlement can also lead to temporal genetic variation in the allele frequencies of a population (Johnson and Wernham 1999). In species with long PLDs and high genetic exchange, post-settlement selection can be large enough to impact recruitment (Butler IV and Herrnkind 1997; Palma et al. 1998). Knowledge of post-settlement selection and its relationship to recruitment variation is especially important in commercial species

(Butler IV and Herrnkind 1997). The Southern rock lobster, *Jasus edwardsii*, represents a valuable fishery in Australia and New Zealand. It occupies a wide geographic range, approximately between 110-180°E and 30-50°S, and sub populations are exposed to different temperature regimes along this latitudinal gradient (Hinojosa et al. 2017). Studies assessing population structure of *J. edwardsii* have distinguished large-scale neutral and adaptive differences between populations in Australia and New Zealand (Thomas and Bell 2013; Morgan et al. 2013; Villacorta-Rath et al. 2016). Although the sampling design of those studies did not allow structure in adult populations to be distinguished at a regional level, larval transport simulations suggest that there is genetic exchange between populations in Australia (Bruce et al. 2007).

The reproductive biology and bi-partite life cycle of the Southern rock lobster makes it a model species to investigate chaotic genetic patchiness and post-settlement selection. *J. edwardsii* is a long lived species, with a life span of more than 30 years (Godwin et al. 2011). Adults reach sexual maturity between two and seven years of age, depending on the latitude (Annala et al. 1980) and they reproduce once a year during the austral winter months (George 2005). As is typical for species exhibiting genetic patchiness, *J. edwardsii* is highly fecund, each female can produce up to half a million eggs per year (Green et al. 2009). Shortly after hatching, phyllosoma larvae are transported offshore and can spend between 12-24 months in this phase before metamorphosing into a post-larva, known as puerulus (Booth 1994). Phyllosoma experience up to 98% mortality during their protracted pelagic larval duration (PLD) (Lesser 1978). This is not uncommon in marine invertebrates, and in the case of a sea urchin, although 2% of survivors were found to be sufficient to replenish a population, they carry a smaller representation of the overall genetic diversity of the adult population (Flowers et al. 2002). Once pueruli settle they remain sedentary and

short distance migration is restricted to nocturnal foraging or retreating to protected areas during moulting (George 2005).

Settlement of *J. edwardsii* pueruli in southeastern Australia is highest between the months of July and February and peaks during winter months (Gardner et al. 2001). Puerulus settlement has been monitored in southeastern Australia for the past four decades using permanent puerulus collectors (Linnane et al. 2010a). In South Australia, Tasmania and New Zealand, a puerulus index based on puerulus catches in collectors is highly correlated to future commercial catches 3-7 years after settlement (Gardner et al. 2001; Linnane et al. 2010a). High variability in year-to-year settlement abundance of *J. edwardsii* and the subsequent fluctuating commercial catch rates have been described and attributed not only to variation in adult reproductive output (Kennington et al. 2013b), but to fluctuating environmental factors prior to settlement. Due to the wide geographical range of *J. edwardsii*, the strength of the environmental factors driving settlement abundance differs from region to region (Hinojosa et al. 2017). Although interannual variation in settlement abundance has been widely studied for *J. edwardsii* (Booth 1994; Hinojosa et al. 2017), no study to date has assessed whether year-to-year fluctuations are also observed in the genetic structure and diversity of pueruli and to what extent natural selection during early post-settlement could affect population structure.

We hypothesized that the combined effect of varying reproductive output between regions and interannual changes in oceanic current patterns leading to varying patterns of settlement would produce spatial and temporal genetic heterogeneity in *J. edwardsii* pueruli and early juveniles. The present study investigates the spatio-temporal genetic variation of newly settled *J. edwardsii* pueruli by assessing genetic structure among four year-classes from two sites in Australia approximately 1,000 km apart from each other. We also studied possible drivers of selection after settlement by

assessing the level of population structure between recently settled pueruli and pueruli that have recently moulted into juveniles. The aims of this study were to: (1) determine whether there is interannual variability in the genetic structure and diversity of recently settled individuals; and (2) to explore the strength and interannual variability of the directional selection acting during settlement.

3.3 Materials and methods

Sampling design

Lobster pueruli were caught in crevice collectors (Booth and Tarring 1986) deployed in South Australia and Tasmania. Pueruli settling into two sites in South Australia, Cape Jaffa and Kingston South East, 19 km apart, were pooled for the analysis and from now on are referred to as "Cape Jaffa". Pueruli settling into two sites in Tasmania, Bicheno shallow and Bicheno deep, separated by less than 1 km, were pooled for the analysis and are referred to as "Bicheno" (Fig. 3.1, Table 3.1). Pueruli were collected monthly from 24 collectors at each region from July to November between years 2009 and 2013, which correspond to periods of highest settlement. Collection of pueruli was carried out by a diver placing a fishnet bag around each collector and attaching a rope with a buoy to each bag. Collectors were hauled into a boat, where they were cleaned (Gardner et al. 2001) and pueruli were immediately stored in 90% ethanol.

Recently settled pueruli were used to assess chaotic genetic patchiness at settlement. However, due to color change after alcohol preservation it was not possible to distinguish the first two puerulus stages. Stage 1 pueruli are recently settled individuals with transparent bodies and no visible digestive tract, whereas stage 2 pueruli are those that settled within the last two days and have a digestive track but no body pigmentation (Booth 1979). For the purposes of this study, stages 1

and 2 were considered together as recently settled individuals and denominated “stage 1-2”.

The other group of pueruli considered in the present study was the “post-puerulus” stage. Post-pueruli have undergone one moult, are fully pigmented and where the pleopods have reduced in size as they are no longer used for forward swimming (Booth 1979). The time between stages 1-2 to post-puerulus is between 20-24 days in winter (Booth and Kittaka 1994) and post-pueruli are approximately 18 months old (Booth 1979).

In order to be able to obtain appropriate sample sizes within each year, puerulus settling between July and November at each sampling year were grouped together and denominated as a “puerulus year". Year 2011 exhibited low puerulus catches in collectors, especially at Bicheno. Only 56 pueruli collected from Bicheno and 119 from Cape Jaffa in 2011 and due to the low abundance of stage 1-2 pueruli, year 2011was excluded from the analysis. Full information on puerulus collector catches across the 5-year period is provided in Fig. S1 (Supporting information).

DNA extraction and ddRADseq library preparation

DNA was extracted from tissue from the horns and legs of each individual using a DNeasy Blood and Tissue kit (Qiagen) and DNA concentration was quantified on a Qubit® 2.0 Fluorometer (Life Technologies). DNA integrity was determined through gel electrophoresis to verify high molecular weight DNA in all samples (>1,000 base pairs [bp] or higher).

ddRADseq libraries were prepared following the protocol described in Villacorta-Rath et al. (2016) using a gel size selection of 400-600 bp in order to maximize the overlapping region among libraries. One interlibrary technical replicate and between three and four intralibrary technical replicates were included in each ddRADseq

library to test for batch effects (Mastretta-Yanes et al. 2015). Electrophoretic assays were performed on all ddRADseq libraries by AGRF using an Agilent BioAnalyser (Agilent Technologies) to determine their exact molecular weight range and library concentration. A total of 330 samples were sequenced across 5.5 lanes at the Australian Genome Research Facility (AGRF) on the Illumina HiSeq 2500 platform using a 100 bp single end kit.

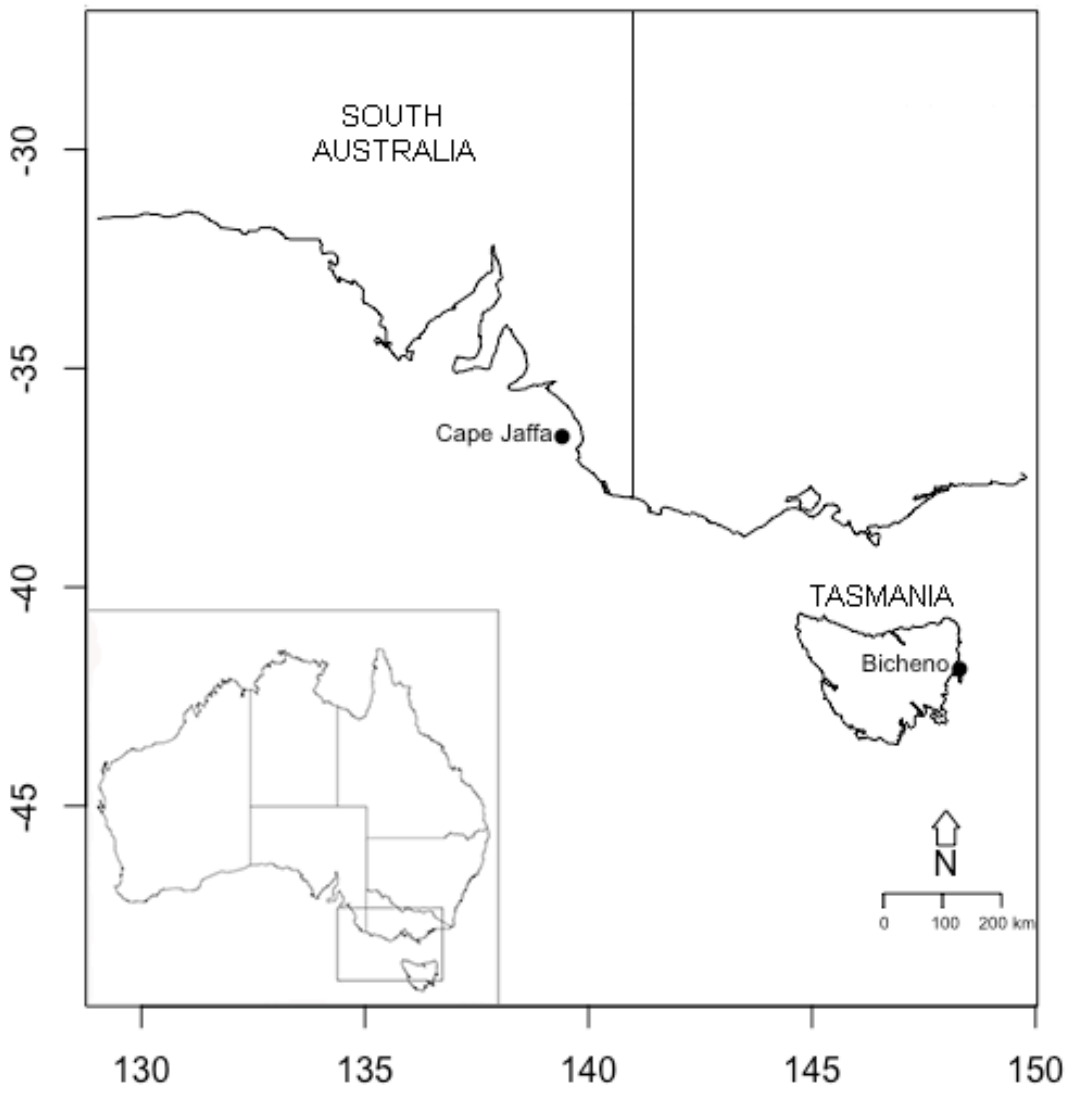


Fig. 3.1 Puerulus collector sites at Cape Jaffa (South Australia) and Bicheno (Tasmania)

Table 3.1 Number of individuals sequenced (n) per year of collection and stage at each collector site

Collector site	State	Latitude	Longitude	Puerulus year	$n_{Stage\ 1\text{-}2}$	$n_{Post\text{-}puerulus}$
Cape Jaffa	South Australia	36°33'S	139°24'E	2009	20	20
				2010	16	18
				2012	19	20
				2013	9	16
Bicheno	Tasmania	41°52'S	148°18'E	2009	18	18
				2010	12	10
				2012	24	15
				2013	20	11

Analyses of raw sequencing data and reference catalog building

An initial quality check of raw indexed data was performed using FastQC v.0.10.1. Data was then demultiplexed using the "process_radtags" protocol from Stacks v.1.29 (Catchen et al. 2011) and hard trimmed to 75 bp (-t 75) to ensure that Phred Quality Score (Q Score) of all reads was above 30. Demultiplexed libraries were filtered for bacterial and viral content using Kraken-gcc v.0.10.4 (Wood and Salzberg 2014).

J. edwardsii possesses a large genome, with a C-value of 5.01 (Deiana et al. 1999) and publicly available pipelines do not perform very well given the amount of repetitive regions contained within it. In our experience, these pipelines result in a reference catalog composed of large amount of paralogous loci. The rad-loci pipeline employed here (and within Villacorta-Rath et al. 2016) was developed in order to cluster raw reads using a series of steps that filter out alleles based on a percentage identity. This process maximizes the removal of paralogous sequences.

Raw reads filtered through Kraken-gcc v.0.10.4 were initially merged and clustered allowing a 4% mismatch (3 bp) between alleles. We allowed each cluster to have a minimum depth of 305 reads. Given that the total number of sequenced samples was 330, this step assumes that most individual samples were represented with at least one read in each cluster. Subsequently, clusters that had less than two and more than 16 different alleles were discarded. The minimum number of two was chosen, as we wanted to obtain bi-allelic data. The maximum number of 16 was based on the fact that for biallelic data, there are 16 possible differences for reads that are divergent in 4 bp. During this step, paralogs and repetitive regions of the genome were discarded. A second round of clustering of the remaining clusters was performed at a 96% identity and a subsequent filtering of clusters that were not composed of a minimum of two and maximum of 16 alleles. Finally, samples were mapped

back to the identified loci and individuals with more than 30% missing data were removed from downstream analyses.

Read alignment, variant calling and neutral SNP characterization

Individual filtered reads were aligned to the reference catalog using bwa-intel v.0.7.12 (Langmead and Salzberg 2012) using the BWA-MEM algorithm. The final alignment was output in the sequence alignment map (SAM) format. SAM files were then converted into their binary form (BAM), sorted and indexed. Subsequently variant calling of the sorted alignments was performed through the Genome Analysis Toolkit (GATK) v.3.3_0 (McKenna et al. 2010) using the Haplotype Caller option. This option outputs a gVCF file with raw, unfiltered SNPs for each sample. All gVCF files were then combined into a single gVCF file. Finally, the gVCF file was transformed into a genotype file (VCF).

A further correction of the VCF file obtained from GATK was performed to ensure the accuracy of the reference and alternate allele calls and to filter out false positives. In the absence of a reference genome, the correction was made by calculating the ratio between the highest quality score over depth ("QD" in VCF file) and lowest QD. If one sample at a specific position had a ratio threshold of 10, which corresponds to a 10% error on a Phred scale, it was substituted by missing data for that SNP.

SNP filtering was performed in vcftools-gcc v.0.1.13 to ensure that only bi-allelic data was present (--min-alleles 2, --max-alleles 2), remove SNPs that were potentially in linkage disequilibrium (--min-r2 0.2), discard SNPs with a minor allele frequency (MAF) of less than 5% (--maf 0.05) and ensure that the minimum SNP depth was 5 (--minDP 5). The maximum amount of missing data for each locus was set to 25% (--max-missing 0.75) and the maximum amount of missing data for each individual was 25%. Individuals with more than

the missing data threshold were removed from subsequent analyses. Finally, only one SNP per locus (--thin 75) was retained.

Neutral loci identification was performed in LOSITAN (Antao et al. 2008) using 100,000 simulations, a confidence interval of 0.95, and a false discovery rate of 0.1 (Jacobsen et al. 2014). LOSITAN uses an F_{ST}-outlier approach that identifies loci are outlier candidates when they exhibit too high or too low F_{ST} compared to neutral expectations (Antao et al. 2008).

In order to minimize false positives in the panel of neutral loci an additional characterization was performed using the R package OutFLANK v.0.1 (Whitlock and Lotterhos 2015). OutFLANK infers the F_{ST} distribution of neutral loci by trimming loci contained in the tails of the distribution, which are likely to be under balancing and positive selection. Therefore the program does not assume a specific population's demographic history and is less prone to false positives than other programs (Whitlock and Lotterhos 2015). The proportion of loci trimmed from both tails of the F_{ST} distribution was set to 5% (LeftTrimFraction = 0.05, RightTrimFraction = 0.05), the minimum heterozygosity required before including calculations from a locus was set to 0 (Hmin = 0) and the false discovery rate was 0.1 (qthreshold = 0.1). Only loci shared between LOSITAN and OutFLANK were considered in the neutral SNPs panel.

Interannual variability in genetic structure of puerulus

To assess interannual variability in genetic structure only recently settled individuals (stage 1-2 pueruli) were analyzed. Global F_{ST} values and confidence intervals for the neutral SNP panel were estimated using the R package mmod v.1.3.2 (Winter 2012). Additionally, pairwise F_{ST} values between sampling sites and years as well as confidence intervals were calculated in the R package hierfstat v.0.04-22 (Goudet 2005). A false discovery rate

correction (FDR) was applied to calculated p-values using the function "p.adjust" of the R package stats v.3.2.4 (R Core Team 2016).

Analyses of Molecular Variance (AMOVA) (Excoffier et al. 1992) were performed in order to quantify the variance explained by (1) puerulus year, (2) sampling site and (3) year*site using the R packages poppr v.2.1.1 (Kamvar et al. 2014) and ade4 v.1.7-4 (Dray and Dufour 2007). Genetic distances were corrected with the "quasieuclidean" method. The number of permutations to test for significance was set to 9999. Finally, discriminant analysis of principal components (DAPC) was used to determine the number of possible genetic clusters based on allele frequency data of both sites across four sampling years. DAPC was performed and results were plotted using the R package adegenet v.1.4-2 (Jombart 2008). The number of clusters present in the dataset was determined using the K-means approach implemented in the function "find.clusters" (Jombart et al. 2010). The optimal number of clusters was the one with the lowest associated BIC (Jombart et al. 2010).

Temporal variation in genetic diversity

Genetic diversity was measured by calculating individual heterozygosity for stage 1-2 puerulus only, in order to obtain the best representation of the genetic diversity in the population before post-settlement mortality could influence diversity. To ensure that heterozygosity of all individuals was measured on the same scale, the standardized heterozygosity per individual (sh) was used. This metric is the proportion of heterozygous loci over the mean heterozygosity across all markers (Coltman et al. 1999). Sh of stage 1-2 puerulus for the neutral SNP panel was calculated using the R package Rhh v.1.0.1 (Alho et al. 2010). To test for differences in the minimum and maximum values of the sh distributions between sampling years, a Mann-Whitney test was performed (Hart 2001). A Bonferroni

correction was applied to the significance level since each sampling year was tested against the other three sampling years and therefore $\alpha = 0.05/3$.

Finally, given that chaotic genetic patchiness can have an effect on the effective population size (Ne), Ne fluctuations were quantified over time at each site. Ne of stage 1-2 pueruli at each sampling year and site was estimated using the NeEstimator software v.2.01 (Do et al. 2014) under the temporal model based on allele frequencies of the neutral SNP panel.

Evidence of natural selection at settlement

Post-settlement selection between stage 1-2 pueruli and post-pueruli was investigated separately for each year and site. To account for potential genetic structure between sites and years an initial outlier SNP identification was carried out for each stage separately as "control" tests and any loci identified here were removed from subsequent analyses. Eight groups were considered for each stage at the control tests (four years and two sites).

SNP characterization was performed using LOSITAN (Antao et al. 2008) and OutFLANK (Whitlock and Lotterhos 2015) with the same set of parameters used for the neutral loci characterization. Only SNPs characterized as being under putative positive selection by both approaches were included in downstream analyses.

The number and magnitude (mean F_{ST} values) of SNPs under putative positive selection was analyzed. Differences in the average magnitude of the selection through time were determined through a Kruskal-Wallis test and post-hoc differences were examined using Mann-Whitney tests. A Bonferroni correction was applied to the significance level at $\alpha = 0.05/5$.

Finally, to determine if the loci containing SNPs putatively under positive selection were contained in protein coding regions, sequences were BLASTed against the complete

Homarus americanus transcriptome (McGrath et al. 2016) and a *J. edwardsii* transcriptome database (SRA Bioproject accession number: PRJNA386609) using BLAST+ v.2.2.29. Queries with statistically significant e-values ($E < 10^{-6}$) and more than 84% identity were considered as valid alignments. Transcriptome sequences that provided significant alignments were annotated using the Trinotate pipeline (https://**trinotate**.github.io/) to determine if they aligned with any known protein domain.

3.4 Results

Sequencing data, SNP filtering and neutral SNP characterization

HiSeq sequencing of 330 pueruli yielded an average of 2.7(±1.03) million reads per sample. A total of 64 samples were discarded due to high percentage of missing data after read alignment, SNP calling and filtering, and 266 samples with sufficient data for downstream analyses remained. The reference catalog built through the rad-loci pipeline produced 5,488 loci. After read alignment, variant calling and filtering for false positive SNPs, we obtained a total of 6,285 variable sites. The final number of SNPs after filtering for MAF, minimum site depth, LD, missing data and retaining one SNP per locus was 900. SNP characterization in LOSITAN identified 558 neutral SNPs, 305 SNPs under putative balancing selection and 37 SNPs under putative positive selection. OutFLANK characterized 888 neutral SNPs, 12 SNPs under putative positive selection and no SNPs under putative balancing selection. The final neutral SNP panel contained 558 loci that were shared between both SNP characterization methods.

A total of nine intralibrary and three interlibrary technical replicates remained after removal of individuals with more than the missing data threshold. A principal components analysis (PCA) was performed to visualize the spatial distribution of replicated samples using the genotypes of all loci before outlier identification (Villacorta-Rath et al. 2016) (Fig. S2,

Supporting information). The PCA plot showed each intralibrary technical replicate pair from both sites distributed close to each other. The three interlibrary technical replicates were also distributed closely in the PCA plot. This suggests that sequencing samples across multiple lanes did not introduce a large bias in the catalog building process.

Interannual variability in genetic structure of recently settled individuals

No genetic differentiation was detected between recently settled pueruli (stage 1-2) of all years and sites when analyzing the neutral SNP panel (global F_{ST} = -0.0042, n.s.) and similarly there was no evidence for population structure found through DAPC (Fig. S3, Supporting information). Although Figure S3 shows two differentiated density distributions, the Bayesian information criteria (BIC) indicated the existence of one cluster and therefore we cannot conclude that there was population structure in the data. However, pairwise F_{ST} values revealed genetic differences amongst most of the collection years in Cape Jaffa, except between years 2009 and 2012 and years 2012 and 2013, and amongst all years in Bicheno (Table 3.2). Significant genetic differences were also observed between Cape Jaffa and Bicheno during most years, except during 2010 and 2013, where no genetic differentiation between sites was found, suggesting a possible common population of origin for those years. A lack of genetic differentiation was also seen between years 2012 in Cape Jaffa and 2009 in Bicheno. The lack of a consistent pattern in genetic differentiation across years suggests that patterns of genetic structure in pueruli are ephemeral as proposed by the chaotic genetic patchiness hypothesis. In support of the differences in pairwise F_{ST} values, allele frequencies of the neutral SNP panel were significantly different between Cape Jaffa and Bicheno, but the degree of differentiation between sites was extremely low (Φ = 0.1%) (AMOVA, Table 3.3). The interaction of sampling site and puerulus year was also significant and explained a small amount of the variation between samples (Φ = 0.2%).

Table 3.2 Nei's pairwise F_{ST} values between stage 1-2 puerulus from Cape Jaffa (CJ) and Bicheno (BIC) across year (2009, 2010, 2012, 2013) using neutral SNPs. Values in bold are significant at an α=0.05 after a FDR correction

	CJ 2009	CJ 2010	CJ 2012	CJ 2013	BIC 2009	BIC 2010	BIC 2012
CJ 2010	**0.0176**						
CJ 2012	0.0028	**0.0195**					
CJ 2013	**0.0073**	**0.0178**	0.0040				
BIC 2009	**0.0094**	**0.0083**	0.0043	**0.0141**			
BIC 2010	**0.0274**	-0.0008	**0.0236**	**0.0240**	**0.0107**		
BIC 2012	**0.0120**	**0.0149**	**0.0062**	**0.0133**	**0.0153**	**0.0175**	
BIC 2013	**0.0094**	**0.0138**	**0.0051**	0.0047	**0.0066**	**0.0139**	**0.0138**

Table 3.3 Results of the AMOVA using genetic distance as a function of: puerulus year, site and the interaction of year and site for the neutral SNP panel

Test	Variance component	Variance	% total	*P*	Φ
Neutral SNPs					
Site	Within samples	91.543	136.139	1	-0.361
	Between samples within site	-24.383	-36.261	1	-0.363
	Between site	0.082	0.123	**0.004**	0.001
Year*Site	Between site within year	0.14	0.209	**0.005**	0.002

Temporal variation in genetic diversity

Genetic diversity of recently settled pueruli estimated using the neutral SNP panel did not differ significantly among years or locations (Fig. 3.2) despite the large year-to-year differences in puerulus catches at both sites (Fig. S1, Supporting information). However, the maximum and minimum values of genetic diversity were variable between years and sites. The minimum values of standardized heterozygosity during years of low puerulus catches were lower than those during years of high puerulus catches at Bicheno but not at Cape Jaffa. This could suggest that years of higher catch rates at Bicheno also exhibited higher genetic

diversity possibly due to the existence of more "winners" contributing to population replenishment at those years.

Finally, Ne estimates for each year as well as the confidence intervals were "infinite" (results not reported) and therefore we could not quantify Ne variation over time and space.

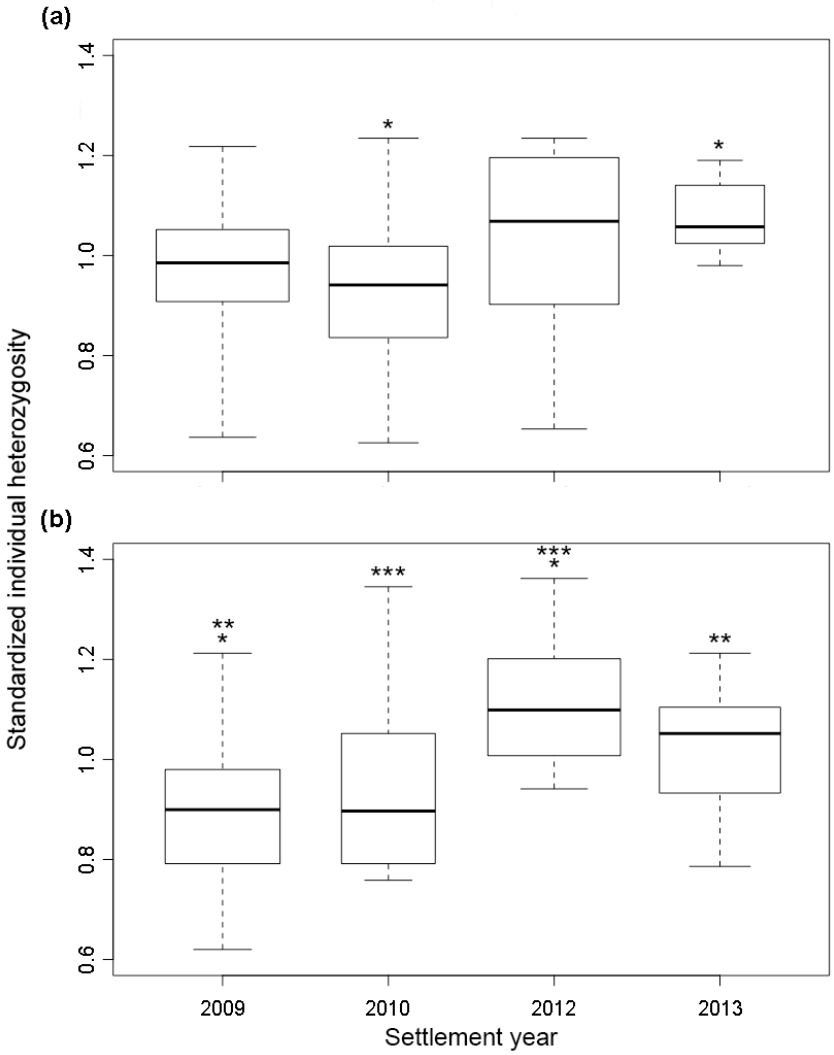


Fig. 3.2 Median standardized individual heterozygosity of stage 1-2 pueruli per settlement year using neutral SNPs for (a) Cape Jaffa and (b) Bicheno. Error bars represent the minimum and maximum observed values of the standardized individual heterozygosity. (*, **, ***) Denote significant differences between years

Genetic identity between sites, developmental stages and between years and the interaction of all factors differed significantly at both sites (Table 3.4). The differences in

allele frequencies between sites, stages and years could be indicative of post-settlement mortality and suggest that this is not constant geographically nor temporally. Pairwise F_{ST} values supported this idea revealing significant differences between developmental stages across almost all years at both sites (Table S1, Supporting information). Post-pueruli from Cape Jaffa collected in 2010 were the exception, as they showed greater genetic similarity to stage 1-2 puerulus from Bicheno in 2009 and Cape Jaffa in 2012.

Table 3.4 Results of the AMOVA using genetic distance as a function of: stage, puerulus year, site and the interaction of stage, year and site for the neutral SNP panel

Test	Variance component	Variance	% total	*P*	Φ
Stage	Within samples	93.985	134.702	1	-0.347
	Between samples within stage	-24.251	-34.757	1	-0.348
	Between stage	0.038	0.055	**0.01**	0.001
Year	Within samples	93.985	134.714	1	-0.347
	Between samples within year	-24.27	-34.788	1	-0.348
	Between year	0.051	0.074	**0.01**	0.001
Site	Within samples	0.065	0.093	1	-0.347
	Between samples within site	-24.264	-34.77	1	-0.348
	Between site	93.985	134.677	**0.01**	0.001
Stage*Year	Between year within stage	0.073	0.105	**0.01**	0.001
Stage*Site	Between site within stage	0.104	0.149	**0.01**	0.001
Year*Site	Between site within year	0.099	0.141	**0.01**	0.001

Evidence of natural selection at settlement

Following the control tests, two loci potentially confounded for selection for year and/or site were identified and removed from the analysis of post-settlement selection. Both sites exhibited a varying number of SNPs under putative positive selection between stage 1-2 and post-pueruli across all years. A single common locus was shared between Cape Jaffa and Bicheno during years 2009 and 2012, supporting common, but weak post-settlement selection over multiple sites/years. All other loci identified as under selection were unique to site and

year. Contrary to predictions of post-settlement selection, genetic diversity did not differ significantly.

Settlement at Cape Jaffa in 2009 exhibited the largest number of outlier SNPs (n = 10) whilst year 2013 exhibited only one outlier locus (Fig. 3.3a). The magnitude of the positive selection at Cape Jaffa also differed significantly across three sampling years (Kruskal-Wallis chi-squared = 8.7829, df = 2, P = 0.01238). Since year 2013 only exhibited one locus under putative positive selection it was excluded from the analysis of variance. Significant differences were detected between years 2009 and 2012 (Fig. 3.3a).

Settlement in year 2012 exhibited the largest number of SNPs under putative positive selection at Bicheno (n = 9), whereas year 2009 exhibited only one SNP under putative positive selection (Fig. 3.3b). The magnitude of the positive selection at this site differed significantly across three sampling years (Kruskal-Wallis chi-squared =8.8676, df = 2, P = 0.01187). Year 2009 was not included in the analysis of variance because there was only one SNP characterized as being under putative positive selection. Therefore the new level of α after the Bonferroni correction was 0.05/3. Post-hoc tests indicated that mean F_{ST} values differed significantly between years 2012 and 2010 (Fig. 3.3b). The level of observed heterozygosity (Ho) remained constant between stage 1-2 and post-pueruli at both sites (Fig. S4, Supporting information), supporting the pattern of weak selection.

Finally, transcriptome sequences from the optical nerve and green gland of *J. edwardsii* exhibited significant hits with three loci containing SNPs under putative positive selection. Additionally, no significant hits between the panel of SNPs under putative positive selection and the *Homarus americanus* transcriptome were found.

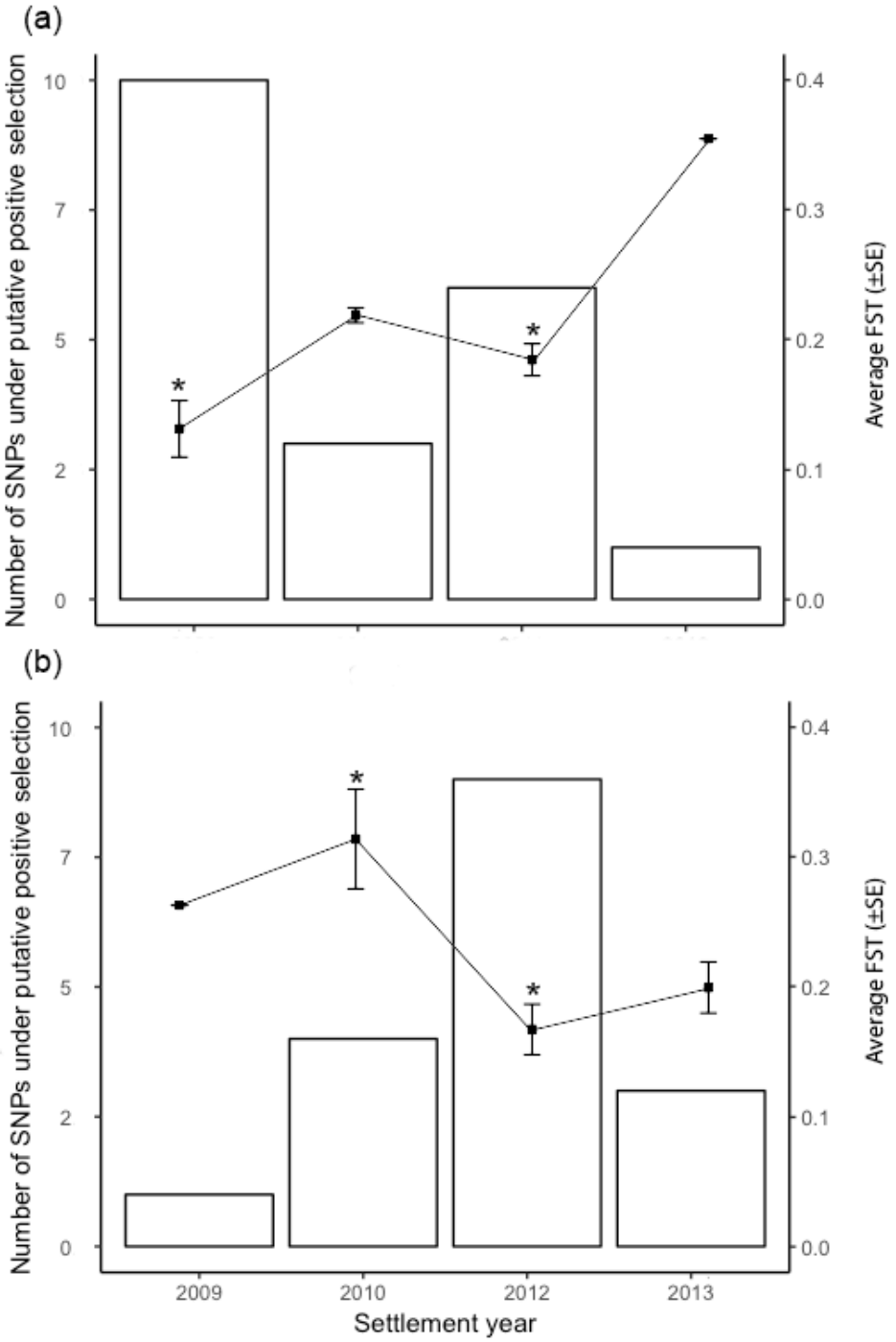


Fig. 3.3 Number of SNPs under putative positive selection and average F_{ST} values per settlement year at (a) Cape Jaffa and (b) Bicheno. The bars represent to the number of SNPs under putative positive selection and the black squares the corresponding average F_{ST} values (±SE). (*) Denotes significant differences in average F_{ST} values between years 2009 and 2012 at Cape Jaffa and years 2010 and 2012 at Bicheno

3.5 Discussion

Our results support chaotic genetic patchiness in *J. edwardsii*. We demonstrated variable patterns of genetic divergence between recently settled individuals in two sampling sites separated by approximately 1,000 km and across four years. Genetic differences in neutral SNPs within and among sites through time as well as differences in the minimum and maximum values of genetic diversity between years of high and low puerulus catches were indicative of chaotic genetic patchiness. There was some evidence of post-settlement selection, which exhibited temporal and spatial variation in the magnitude and strength of selection following settlement.

Chaotic genetic patchiness

The significant genetic differentiation between stage 1-2 pueruli across sampling sites and years found in the present study point to the existence of chaotic genetic patchiness. Genetic patchiness is a widespread phenomenon occurring in marine invertebrates, including spiny lobsters (Thompson et al. 1996; Johnson and Wernham 1999; Iacchei et al. 2013; Kennington et al. 2013a; Truelove et al. 2017). For example, temporal variation in allele frequencies of two allozymes revealed ephemeral structure in the adult population of the western rock lobster, *Panulirus cygnus* (Thompson et al. 1996). Since *P. cygnus* constitutes a single panmictic stock along its whole distribution, the authors attributed the temporal discrepancies to variation in recruitment among the study sites (Thompson et al. 1996). Similarly, unexpected patterns of genetic heterogeneity were found the adult population of the California spiny lobster, *Panulirus interruptus*, despite no evidence of isolation-by-distance throughout its geographic range (Iacchei et al. 2013). This genetic structure was attributed to localized recruitment of related individuals (Iacchei et al. 2013).

The significant interaction between site and year (AMOVA neutral SNPs, Table 3.3) found in the present study and the significantly different F_{ST} values between years within both sites found herein further support the existence of chaotic genetic patchiness. In the case of *J. edwardsii*, while examination of the adult populations generally concludes panmixia within Australia (Thomas and Bell 2013; Morgan et al. 2013; Villacorta-Rath et al. 2016), our analysis including cohorts of pueruli over time and space blurs this conclusion. Recently settled individuals (stage 1-2 pueruli) analyzed in the present study exhibited a non-significant global level of genetic differentiation across sites and years (F_{ST} = -0.0042), suggesting that there is gene flow between South Australia and Tasmania. However pairwise F_{ST} values differed significantly both geographically and temporally. This incongruence is indicative of genetic patchiness. Similar trends in lack of global population differentiation but fine-scale genetic heterogeneity has been reported in *P. cygnus* (Kennington et al., 2013a). The authors hypothesized that even in a panmictic population, such as that of *P. cygnus*, genetic heterogeneity can arise from sweepstakes reproductive success (SRS) or pre-settlement mortality (Kennington et al. 2013a).

Studies assessing chaotic genetic patchiness during early life stages of lobsters have also identified ephemeral genetic structure. Johnson & Wernham (1999) sampled recently settled *P. cygnus* in two sites 350 km apart during three consecutive recruitment seasons. Allele frequencies of one allozyme marker differed significantly between sites only during one recruitment season. Given that the pueruli settling into each of the two sites at a given recruitment season were genetically homogenous, the authors concluded that SRS was not a driver of genetic patchiness. In the present study, pueruli also showed genetic homogeneity within sites at each year (Figs. S5, S6, Supporting information), suggesting that the observed ephemeral population structure is possibly not a result of differential reproduction. However, the significant differences observed in the minimum and maximum values of the standardized

individual heterozygosity between years of high and low puerulus catches at Bicheno could be due to the existence of a higher number of genetically different groups contributing to the next generation of settlers.

Estimates of Ne and the Ne/N ratio across years could help determine the cause of chaotic genetic patchiness. One possible cause of genetic patchiness in recently settled individuals is differential adult reproduction (Johnson and Black 1982). In the presence of SRS, Ne is much smaller than the census size of a population, given that a minority of individuals reproduces successfully (Hedgecock 1994). Therefore large variations in Ne through time could be indicative of SRS. Unfortunately, we were unable to determine if effective population size fluctuates through time due to the infinite values of Ne obtained, probably due to the low sample size per site and year. However, given the lack of significant differences in the median levels of genetic diversity found across the four-year sampling period of the present study, it is unlikely that we would have been able to distinguish differences in Ne between populations in the presence of larger sample sizes.

Ocean advection and retention of larvae can generate unexpected genetic patchiness in highly dispersive organisms. A biophysical model of larval dispersal explained a large percentage of the variation in population structure between populations of *Panulirus argus* across the Caribbean (Truelove et al. 2016). The authors identified an oceanographic boundary dividing the region into two provinces: a southern province characterized by recirculation and a northern province influenced by offshore currents. These two distinct oceanic features created genetic patchiness along *P. argus'* geographic range (Truelove et al. 2016). In the case of *J. edwardsii*, patterns of settlement abundance have also been attributed to oceanographic features that vary in strength and importance among regions (Hinojosa et al. 2017). An oceanographic model simulating larval transport of *J. edwardsii* suggested that there is a high degree of self-recruitment in areas such as Cape Jaffa (Bruce et al. 2007).

Conversely, East Tasmania (Bicheno) receives settlers that originated from South Australia and Tasmania but that fluctuations in ocean advection create year-to-year variability in the source of pueruli (Bruce et al. 2007). In the present study, the different environmental factors influencing larval transport in South Australia and Tasmania could explain the clear structure between sites whereas the year-to-year variability in oceanic features could be behind the genetic structure between years at each site.

Post-settlement selection

We identified some evidence of weak positive selection acting shortly after settlement of *J. edwardsii* that is consistent across site and year, however the majority of the results suggest variable selection through space and time. Predation during early post-settlement of spiny lobsters is extremely high and is considered a potential bottleneck to recruitment (Smith and Herrnkind 1992). At the same time, predation risk in recently settled clawed and spiny lobsters is highly dependent on shelter availability (Butler IV et al. 1997). For example, survival of the Caribbean spiny lobster during the "algal-phase" was positively related to settlement habitat suitability (Butler IV et al. 1997). The common locus between sites and years found herein could be linked to a trait that makes individuals more vulnerable to predation or affect their ability to settle into suitable habitat. However, we should interpret these results with caution since there was only one SNP under putative positive selection shared between both sites and we were not able to ascribe a function to this SNP.

The majority of evidence for post settlement selection results from different panels of putative positively selected SNPs identified for each year and site. This suggests that either the selective forces acting during early post-settlement of *J. edwardsii* differ across the geographic range of the species and from year to year, or that different loci are responding to the same selection pressure. Differences in selection pressures would be the most likely

scenario given that climate-ocean forces driving settlement in *J. edwardsii* are also variable across the species distribution (Hinojosa et al. 2017). As mentioned above, predation is the main source of mortality after settlement. Since the vulnerability to post-settlement predation decreases with individual size (Butler IV and Herrnkind 1997; Palma et al. 1998), variability in factors affecting growth across sampling sites and years may play a role in the divergence in magnitude and strength of the selection found herein. Water temperature has been described as the main abiotic factor determining intermoult interval and moult increments in the early benthic phase of Caribbean spiny lobsters (Forcucci et al. 1994). Although that study sampled only one site, lobster growth was determined during summer and winter months. The temperature gradient between seasons in Forcucci et al.'s study was of approximately 10°C, which is higher than the temperature gradient between sampling sites of the present study. However, given that *J. edwardsii* exhibits high phenotypic plasticity, slight variation in water temperature can have large impact on growth and size-at-age (Chandrapavan et al. 2010).

Processes occurring prior to settlement can also influence post-settlement mortality in marine invertebrates (Palma et al. 1998). For example, inadequate food intake during the larval phase can reduce juvenile growth rates (Pechenik and Tyrell 2015). Since plankton abundance fluctuates spatially and temporarily (Jeffs et al. 2004; Wang et al. 2015), condition of *J. edwardsii* phyllosoma could potentially be very variable throughout its prolonged PLD. Latent effects of larval development on subsequent stages have been reported for a wide range of invertebrates (Pechenik 2006). Consequently, phyllosoma that developed under sub-optimal feeding conditions could experience reduced growth after settlement and therefore be more vulnerable to mortality during this period (Pechenik 2006).

Interestingly, the years of highest puerulus catches at both sites also exhibited the largest number of SNPs under putative positive selection. This could suggest that *J. edwardsii*

experiences density-dependent mortality at settlement, however this hypothesis is highly unlikely. Recently settled spiny lobster pueruli exhibit solitary living in macroalgae (Marx and Herrnkind 1985) however their survival and growth is influenced by habitat complexity rather than conspecific density (Butler IV et al. 1997). The existence of more loci under putative positive selection during years of higher puerulus catches at each site (year 2009 at Cape Jaffa and year 2012 at Bicheno) could be due to the larger sample sizes available for those years. Nevertheless, years 2009 and 2012 at both sites had very similar sample sizes and yet there were large differences in the number of loci under putative positive selection. Therefore the variability in the loci could indicate that there is varying selection pressure among years of higher and lower puerulus catches.

Finally, the constant pattern of observed heterozygosity between stage 1-2 and post-pueruli could mean that neither homozygote nor heterozygote genotypes are favored by the weak selection found herein. Another possible explanation for the lack of differences in Ho among developmental stages is that the post-pueruli analyzed in the present study do not comprise the same genetic cohort as recently settled pueruli. If multi-cohorts were sampled (stage 1-2 pueruli being genetically different cohort from post-pueruli), comparisons of genetic diversity between stages would give biased results. However, the sampling design of this study prevents us from being able to answer this question and therefore a fine-scale study assessing how the genetic identity of settlers arriving in consecutive pulses within a settlement season fluctuates could clarify this subject.

3.6 Conclusions

The significant genetic differences at neutral loci between some of the sampling years were consistent with the existence of chaotic genetic patchiness in *J. edwardsii*. In the presence of high self-recruitment and high egg production in South Australia, suggested by

an earlier oceanographic model, population structure at the northernmost site could be mainly driven by ocean advection influencing the number of settlers arriving at this site every year. At the southernmost site, structure could be primarily driven by fluctuations in the amount of source populations producing successful settlers every year.

A weak positive post-settlement selection was identified, however, selective forces acting during early post-settlement of *J. edwardsii* differ by site and time, providing further evidence of chaotic genetic patchiness. The possibility of having sampled multiple cohorts could also be driving the observed pattern of fluctuating positive selection between developmental stages. In the former case, differential mortality across a geographic range could be due to environmental factors and also genotype-related fitness of recently settled individuals.

3.7 Acknowledgements

Funding for this research was provided by an Australian Research Council Linkage Project grant (Project No. LP120200164) and Australian Research Council Discovery Project grant (Project No. DP150101491). We would like to thank Peter Hawthorne (South Australian Research and Development Institute, SARDI) for puerulus collections and data on settlement abundance from South Australia. We thank Ivan Hinojosa, Ruari Colquhoun, David Fallon, Kylie Cahill, Graeme Ewing and Sarah Pyke for puerulus collections in Tasmania. Thanks to Brooke Whitelaw and Daniel Laird for laboratory assistance. Special thanks to Karen J. Miller who contributed to the original project idea.

3.8 Supplementary information

Table S1 Nei's pairwise F_{ST} values between stage 1-2 and post-puerulus from Cape Jaffa (CJ) and Bicheno (BIC) using neutral SNPs. Values in bold are significant at an α=0.05

		POST-PUERULUS							
		CJ2009	CJ2010	CJ2012	CJ2013	BIC2009	BIC2010	BIC2012	BIC2013
STAGE 1-2	CJ2009	**0.0056**	**0.0036**	**0.0162**	**0.0135**	**0.0160**	**0.0158**	**0.0107**	**0.0149**
	CJ2010	**0.0255**	**0.0101**	**0.0256**	**0.019**	**0.0287**	**0.0177**	**0.0223**	**0.0103**
	CJ2012	**0.0072**	0.0035	**0.011**	**0.0109**	**0.0115**	**0.0144**	**0.0118**	**0.0116**
	CJ2013	**0.0101**	**0.0070**	**0.0062**	**0.0079**	**0.0091**	**0.0110**	**0.0119**	**0.0102**
	BIC2009	**0.0156**	0.0024	**0.0175**	**0.011**	**0.0200**	**0.0114**	**0.0069**	**0.0065**
	BIC2010	**0.0339**	**0.0143**	**0.0267**	**0.0192**	**0.0307**	**0.0219**	**0.0272**	**0.0073**
	BIC2012	**0.0124**	**0.0109**	**0.018**	**0.0178**	**0.0138**	**0.0179**	**0.0206**	**0.0180**
	BIC2013	**0.0142**	**0.0065**	**0.0106**	**0.0069**	**0.0148**	**0.0132**	**0.0072**	**0.0047**

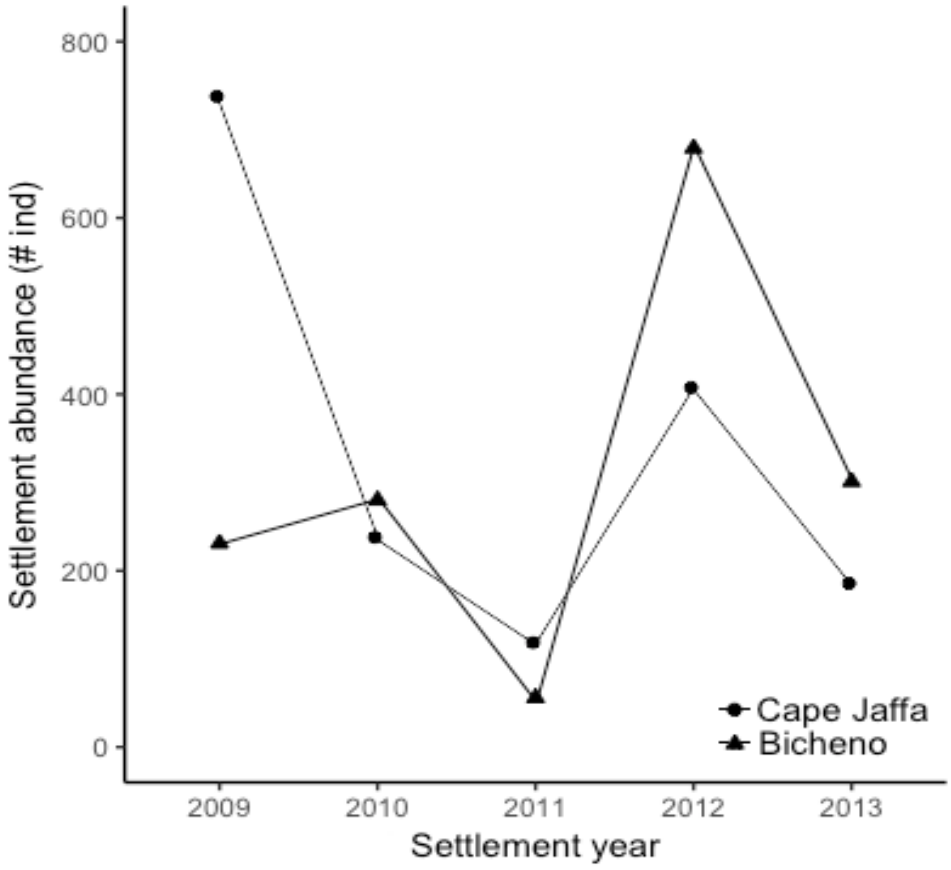


Fig. S1 Puerulus collector catches per site throughout a five-year period. Numbers of puerulus represent the total puerulus catch in collectors, including stages 1, 2, 3 and post puerulus

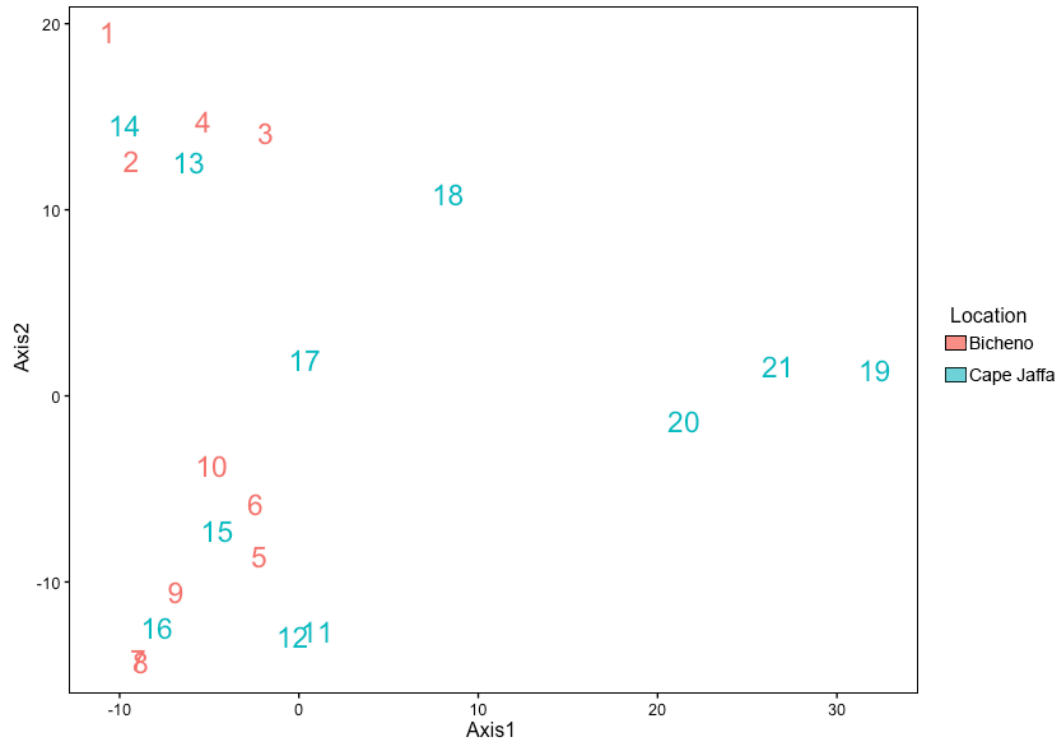


Fig. S2 Principal component analysis based on allele frequencies of nine intralibrary replicate pairs from Bicheno (red numbers) and Cape Jaffa (blue numbers). Consecutive numbers from 1 to 18 belong to each intralibrary technical replicate pair (e.g. 1&2, 3&4, etc. represent replicate pair). Samples 19, 20 and 21 constitute interlibrary technical replicates

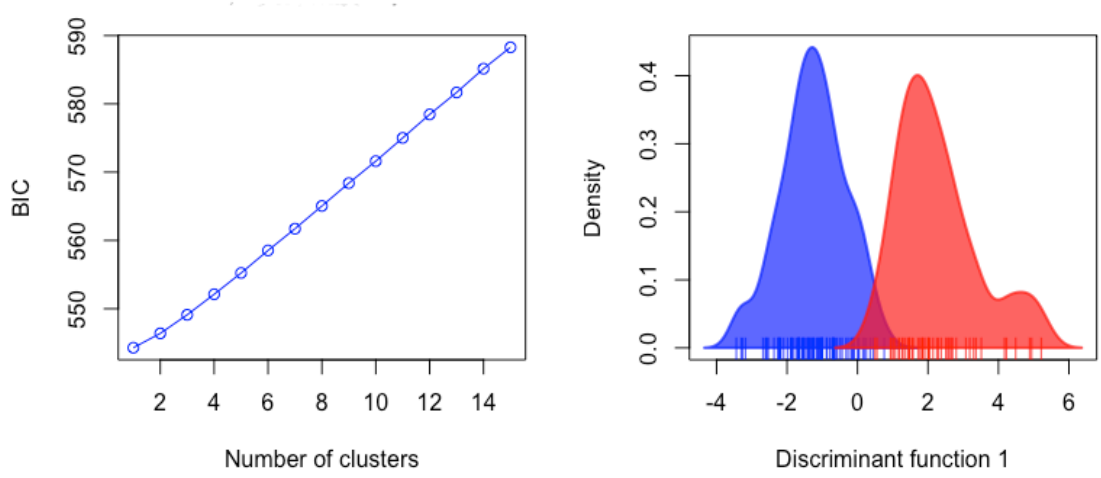


Fig. S3 Genetic differentiation between stage 1-2 pueruli between Bicheno and Cape Jaffa across sampling years using neutral SNPs. The left panel shows the Bayesian index criteria (BIC) plots used to define the most likely number of groups present. The right panel shows the first principal component resulting from the discriminant analysis of principal components

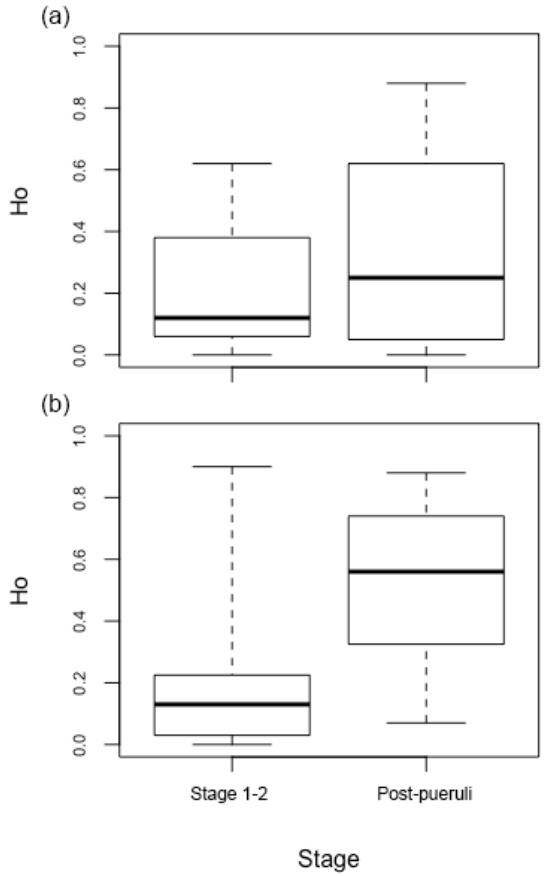


Fig. S4 Median observed heterozygosity per stage, pooled across years, using SNPs under putative positive selection for (a) Cape Jaffa and (b) Bicheno. Error bars represent the minimum and maximum observed values of heterozygosity

Supplemental Materials and methods

Cohort identification in recently settled pueruli

In order to identify cohorts of stage 1-2 pueruli arriving at the same location during each year a discriminant analysis of principal components (DAPC) was performed using the R package adegenet v.1.4-2 (Jombart 2008). Eight DAPC analyses were carried out to determine the number of possible genetic clusters based on allele frequency data for each year and each site. In each case, the maximum number of clusters (max.n.clust) was set to half of the sample size. The most likely number of clusters was chosen using the K-means approach implemented in the function "find.clusters" and given by the cluster with the lowest associated Bayesian index criteria (BIC) (Jombart et al. 2010).

Supplemental Results

Cohort identification in recently settled pueruli

Lack of genetic differentiation between stage 1-2 pueruli suggested the absence of cohorts of individuals settling during the same year into each site (Figs. S5, S6). Although all DAPCs exhibited two well-separated density distributions, the minimum BIC denoted the existence of only one genetic group.

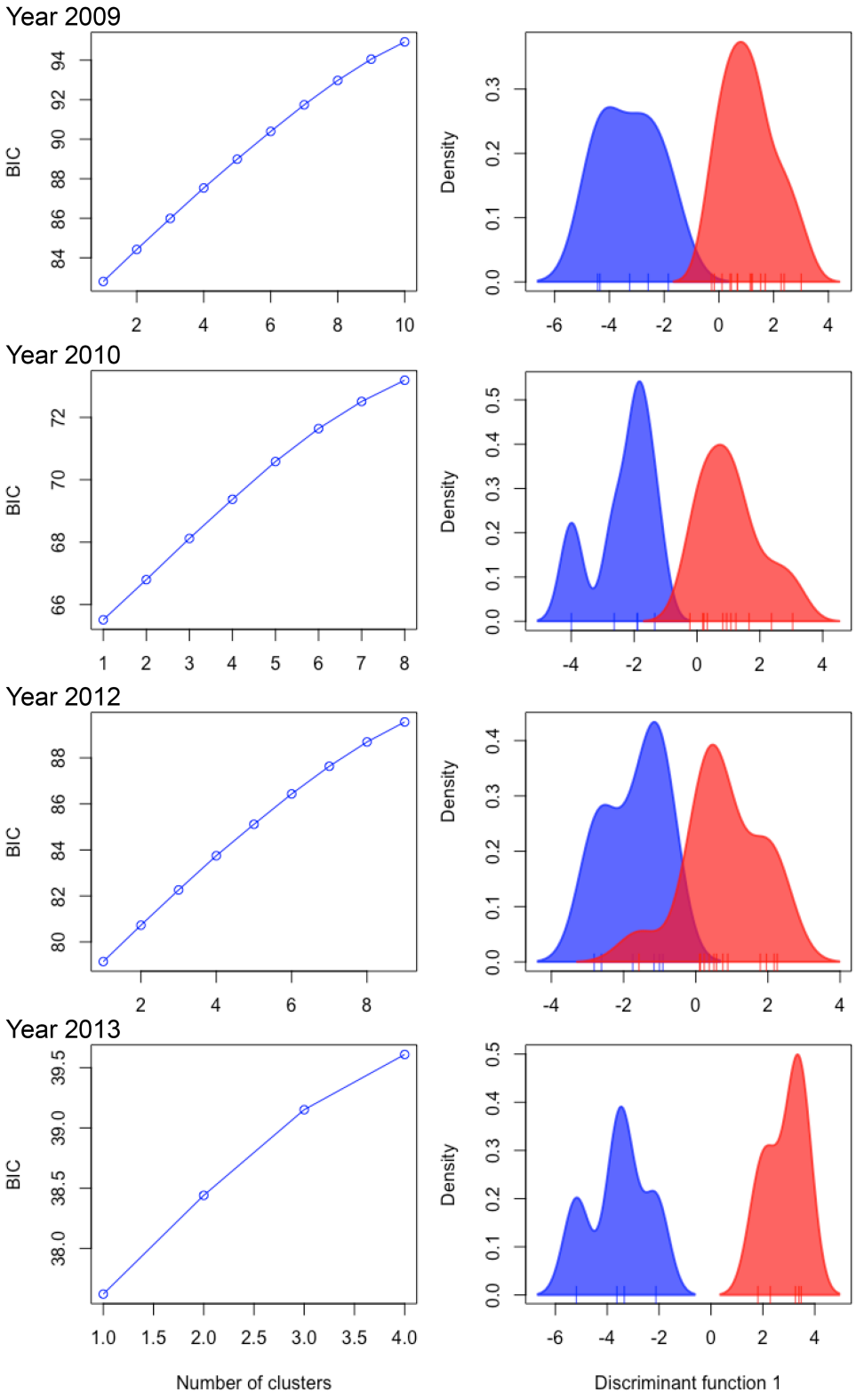


Fig. S5 Genetic differentiation between stage 1-2 pueruli from each sampling year at Cape Jaffa using neutral SNPs. The left panels show the Bayesian index criteria (BIC) plots used to define the most likely number of groups present. The right panels show the first principal component resulting from the discriminant analysis of principal components

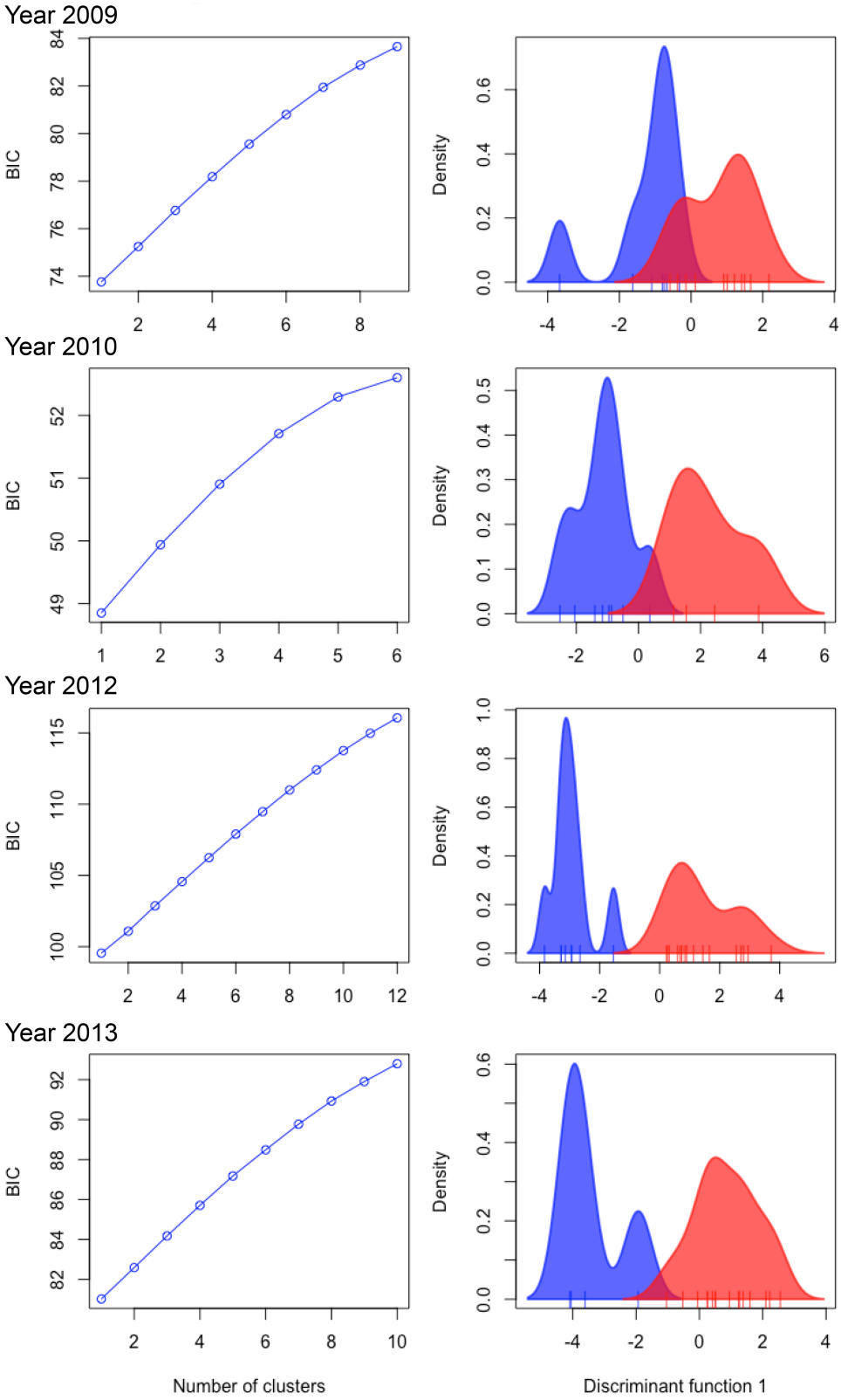


Fig. S6 Genetic differentiation between stage 1-2 pueruli from each sampling year at Bicheno using neutral SNPs. The left panels show the Bayesian index criteria (BIC) plots used to define the most likely number of groups present. The right panels show the first principal component resulting from the discriminant analysis of principal components

Chapter 4: Can dispersal history drive spatio-temporal phenotypic differences between Southern rock lobster (*Jasus edwardsii*) post-larvae?

4.1 Abstract

Recruitment of marine organisms with bi-partite life cycle is the result of processes taking place from embryogenesis to settlement. Most importantly, environmental conditions experienced during dispersal and settlement can determine recruitment success. This is particularly interesting to detect in species that exhibit seasonal reproduction, whereby slight differences in egg extrusion times can produce large variability in the dispersal fate of consecutive cohorts. When egg extrusion timing has a genetic reason (i.e. isolation by time), phenotypic differences between cohorts of recruits become even larger. Therefore phenotypic variability can arise in recruits due to the different genotypes and/or dispersal histories. Here we investigate whether morphological differences in recently settled Southern rock lobster (*Jasus edwardsii*) recruits along a latitudinal and temporal gradient in Tasmania, Australia, are due to natural selection or to dispersal history. We used double digest restriction-site associated DNA sequencing (ddRADseq) to assess differences in genetic structure of recently settled recruits on the east coast of Tasmania over three months of peak settlement during 2012. Phenotypic differences in pueruli between sites and months of settlement were observed, with smaller individuals found at the northernmost site. There was a lack of overall genetic divergence, but significant differences in pairwise F_{ST} between settlement months at the southernmost study site suggested chaotic genetic patchiness. Individuals settling into that site earlier in the season were genetically different from those settling later, providing evidence for isolation by time at this site. These results suggest that dispersal history can drive phenotypic divergence in *J. edwardsii* recruits.

4.2 Introduction

In marine organisms with a planktonic larval phase & benthic adult phase, recruitment, or the passage from pelagic to demersal life, is temporally and spatially variable (Sale 1990). This is because it is highly dependent on biological factors that determine larval supply and quality as well as on the environmental factors experienced during dispersal and settlement (Rodriguez et al. 1993; Gimenez 2006). This variability does not only encompass fluctuations in the abundance of individuals recruiting into the population, but also refers to phenotypic differences among recruits (Moland et al. 2010; Butler et al. 2015). Maternal fecundity, as well as energy investment on the offspring, is an important factor controlling larval supply and quality (Green et al. 2014). Among the environmental factors accounting for variability in abundance and phenotype of recruits, water temperature and food availability are the most studied, as they influence growth and therefore time to metamorphosis and size-at-settlement (Cobb et al. 1997). Finally, settlement habitat represents a source of recruitment variability itself since recruits must be adapted to the specific settlement environment (Cobb et al. 1997).

Phenotypic plasticity of recruits is highly dependent on maternal contributions to the offspring (Green 2008). In general female fecundity increases with body size, given that mothers can allocate more energy reserves to the eggs (Moland et al. 2010). A laboratory experiment testing maternal effects on *Homarus gammarus* eggs and larvae showed that larger females consistently produced larger offspring with increased pelagic larval survival (Moland et al. 2010). Additionally, larger offspring may have increased chances of survival because they can endure suboptimal environmental conditions, namely food availability (Suzuki and Sato 2010). For example, larger larvae of the Kuruma prawn, *Marsupenaeus japonicus*, hatching from larger females were able to resist periods of starvation (Sato et al. 2017). However, larger *M. japonicus* larvae exhibited a survival advantage at recruitment over smaller larvae only in the presence of food limitation (Sato et al. 2017). Similarly,

development of the larval marine shrimp, *Palaemon serratus*, was influenced by the environment and size-at-hatching, which was determined by maternal body size (González-Ortegón and Giménez 2014). This evidence suggests that maternal influences on the offspring are ultimately regulated by the environment experienced during early life stages (Marshall and Uller 2007).

The environment can exert influence on the recruits' phenotype by acting on genetic factors, such as local adaptation, or non-genetic factors, such as adaptive plasticity (Marshall et al. 2010). Local adaptation ensures the persistence of certain phenotypes in a population (Schmid and Guillaume 2017). If migration and natural selection are balanced, the cline in phenotypic variability across space would be gradual and phenotypic differences would also correspond to differences in allele frequencies between populations (Schmid and Guillaume 2017). In contrast, when the environment is the sole driver of phenotypic differentiation, phenotypic differences in the population would not match the observed genetic divergence (Schmid and Guillaume 2017). In this case, the environment constitutes a selective pressure during larval dispersal and settlement that can be translated into genetic patchiness (Eldon et al. 2016).

The non-genetic phenotypic differences arise because the environment has profound effects on the metabolism and growth of larvae (Marshall et al. 2003). Moreover, its influence during early life can be carried over into subsequent stages (Marshall et al. 2003). Cohorts of marine invertebrate recruits exhibit high phenotypic plasticity due to the wide range of environmental conditions experienced through the pelagic larval duration (PLD) (Rey et al. 2016). For example, if egg hatching periods coincide with low food availability, larvae may experience poor survival or high post-settlement mortality (Gimenez 2006). If dispersal history determines the phenotype of recruits, cohorts of larvae hatching at different times during a season could thus exhibit large phenotypic differences at settlement. It is then

possible that females becoming reproductive at different times of the season would result in recruits with slightly different phenotypes (Kunisch and Anger 1984; Gimenez 2006). Although there is evidence suggesting that there are differences in the size of recruits of invertebrates across a geographic and temporal scale, it is unknown whether these differences are purely driven by dispersal history or if local adaptation can also play an important role in shaping recruitment.

The Southern rock lobster, *Jasus edwardsii* is an ideal model species to study the link between phenotype and genotype due to its large reproductive output and protracted pelagic larval duration. This species is highly fecund, producing approximately 500,000 eggs per brood (Green et al. 2014). Adults experience seasonal reproduction that extends from May to September but egg-hatching periods vary slightly across populations through a latitudinal gradient (Powell et al. 2016). *J. edwardsii* exhibits the largest PLD of all species within the palinurid family, spending between 12 and 24 months before metamorphosing into a puerulus post-larva (Booth 1994). This extended time gives a large window for environmental factors to shape larval phenotypes. Recruitment monitoring of this species in southeast Australia has evidenced large variability in abundance (Linnane et al. 2014a; Hinojosa et al. 2017) as well as in size-at-settlement (Booth 1979, 1994), as it is expected for highly fecund species (Cobb et al. 1997).

Based on previous evidence of phenotypic differences in recruits, we hypothesized that dispersal history would impact on the phenotype and genetic identity of recently settled individuals. The present study explores patterns of morphological and genetic differences in recently settled *J. edwardsii* along a latitudinal and temporal gradient in Tasmania, Australia. More specifically, the aims of this study were: (1) to compare size-at-settlement of *J. edwardsii* pueruli at three sampling sites during three consecutive settlement months, (2), to investigate if phenotypic plasticity was related to neutral and non-neutral genotypic variation

and (3) to determine whether there was a pattern of isolation by time (IBT) in recently settled *J. edwardsii*.

4.3 Materials and methods

Sample collection

Lobster pueruli were caught in crevice collectors (Booth and Tarring 1986) deployed at three sites on the east coast of Tasmania, southeast Australia (Fig. 4.1) (Gardner et al. 2001). Collectors were attached to permanent structures fixed on sandy substratum at three to nine metres depth (Gardner et al. 2001). Puerulus collectors were serviced monthly by a diver who placed a mesh bag around each collector and attached a rope with a buoy to each bag. Collectors were then hauled into the boat, where they were cleaned (Gardner et al. 2001). All pueruli were removed from the collectors and immediately stored in 90% ethanol. For the purpose of this study, recently settled pueruli (stages 1 and 2) during the months of August, September and October 2012 were considered (Table 4.1).

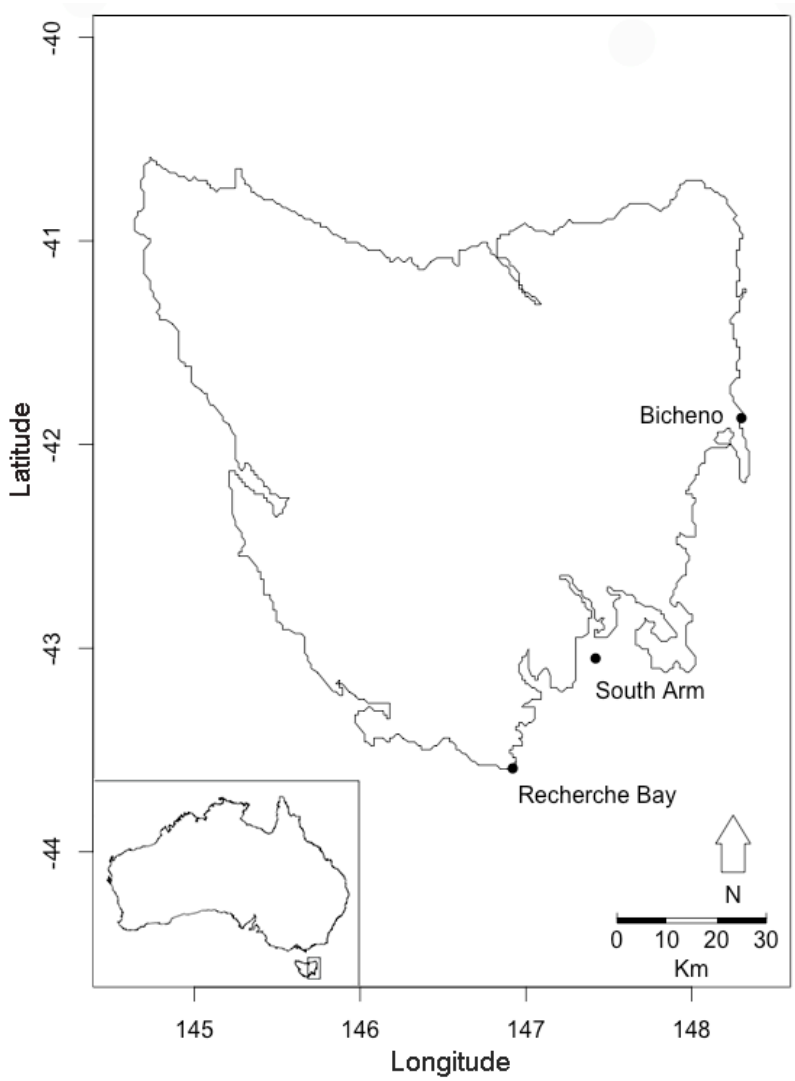


Fig. 4.1 Puerulus collector sites in Tasmania, Australia

Table 4.1 Number of individuals analyzed per month of collection during the year 2012 at each collector site. Number of individuals used for morphometric analyses (N_{morph}) and number of individuals used for sequencing (N_{seq}) are reported

Collector site	Latitude	Longitude	Month of collection	N_{morph}	N_{seq}
Bicheno	41°52'S	148°18'E	August	19	14
			September	17	7
			October	15	15
South Arm	43°03'S	147°25'E	August	20	16
			September	12	10
			October	3	4
Recherche Bay	43°35'S	146°55'E	August	15	11
			September	13	12
			October	13	15

Comparing morphometrics of pueruli

Pueruli individuals were removed from ethanol, blot dried to eliminate the excess ethanol and weighed (g ± 0.01). To determine whether pueruli weight varied with settlement month, site and the interaction of both a two-way analysis of variance (ANOVA) was performed using the function "aov" implemented within the R package stats v.3.2.4 (R Core Team 2016). A post-hoc examination of differences between individual weights was carried out using the Tukey's "Honest Significance Difference" method using the function "TukeyHSD" of the R package stats v.3.2.4.

DNA extraction and ddRADseq library preparation

DNA was extracted from tissue from the horns and legs of each individual using the DNeasy Blood and Tissue kit (Qiagen) following the manufacturer's instructions. DNA concentration was quantified on a Qubit® 2.0 Fluorometer (Life Technologies). DNA integrity was determined through gel electrophoresis to verify that high molecular weight DNA was obtained in all samples (1,000 base pairs [bp] or higher).

Four ddRADseq libraries were prepared using a modified version of (Peterson et al. 2012) ddRAD protocol, described in Villacorta-Rath et al. (2016). Gel size selection was set to 400-600 bp in order to maximize the overlapping region among all ddRAD libraries and therefore increase the chances of encountering common loci. A total of 59-60 samples were pooled in each library. Additionally, within each library 9-10 technical replicates were included in order to test for sequencing-derived differences between libraries (Mastretta-Yanes et al. 2015). Electrophoretic assays of pooled libraries were run using an Agilent BioAnalyzer (Agilent Technologies) to determine their exact molecular weight range and library concentration. Pooled libraries were sequenced at the Australian Genome Research Facility (AGRF) on the Illumina HiSeq 2500 platform using a 100 bp single end kit. After demultiplexing, 28 samples that yielded a lower than average number of reads were re-sequenced in half a lane of the Illumina HiSeq 2500 platform using a 100 bp single end kit.

Analyses of raw sequencing data and reference catalog building

An initial quality check of raw indexed data was performed using FastQC v.0.10.1. Data was then demultiplexed using the "process_radtags" protocol from Stacks v.1.29 (Catchen et al. 2011) and hard trimmed to 75 bp to ensure that Phred Quality Score (Q Score) of all reads were above 30. Finally, demultiplexed libraries were filtered for bacterial and viral content using Kraken-gcc v.0.10.4 (Wood and Salzberg 2014).

Filtered reads were processed through a custom-built "rad-loci" pipeline (https://github.com/molecularbiodiversity/rad-loci) to obtain a catalog of reference loci detailed within Villacorta-Rath et al. (2016). An initial run of the rad-loci pipeline, read alignment and SNP calling was performed and samples with more than 25% missing data were discarded in order to maximize the final number of polymorphic loci obtained. Due to a high percentage of missing data, only a subset of the individuals whose morphometrics were

measured were included in the genetics analyses. A total of 70 samples were discarded and 193 samples with less than 25% missing data remained. The subsequent rad-loci pipeline run was therefore performed with a subset of 193 samples using VSearch v.1.1.3 (Rognes et al. 2016). Initially, all reads were pooled and clusters with a depth of at least 193 were retained, assuming that there would be at least one read per sample represented in each cluster. Subsequently all clusters that were comprised of reads with more than a 4% mismatch (3 bp) were discarded. Assuming that each member of a cluster was an allele, only clusters that had between two and 16 members were retained. Finally, only samples with a minimum of 10,000 alleles and 30% missing data were retained for subsequent read alignment and SNP calling.

Read alignment, variant calling and SNP filtering

Individual filtered reads of all samples were aligned to the reference catalog using bwa-intel v.0.7.12 (Langmead and Salzberg 2012). Variant calling was performed using the Genome Analysis Toolkit (GATK) v.3.3_0 (McKenna et al. 2010), with a further correction of the variant call format (vcf) file performed to ensure the accuracy of the reference and alternate allele calls, and to filter out false positives. In the absence of a reference genome, the correction was made by calculating the ratio between the highest quality score over depth ("QD" in vcf file) and lowest QD. If one sample at a specific position had a ratio threshold of 10, which corresponds to a 10% error on a Phred scale, the SNP at that position was replaced with missing data.

SNP filtering was performed in vcftools-gcc v.0.1.13 to ensure that only bi-allelic data was present (--min-alleles 2, --max-alleles 2), to remove SNPs that were potentially in linkage disequilibrium (--min-r2 0.2), to discard SNPs with a minor allele frequency (MAF) of less than 5% (--maf 0.05), and to ensure that the minimum SNP depth was 5 (--minDP 5).

The maximum amount of missing data for each locus was set to 25% (--max-missing 0.75) and individuals with more than 25% missing data were removed from subsequent analyses. Finally, only one SNP per locus (--thin 75) was retained.

A total of 22 technical replicates remained after removal of individuals with more than the missing data threshold. A principal components analysis (PCA) was performed to visualize the spatial distribution of replicated samples.

Outlier SNP identification

In order to minimize false positives in the panel of SNPs under putative positive selection, two outlier identification methods were used. The first SNP characterization was performed in the software LOSITAN (Antao et al. 2008) using 100,000 simulations, a confidence interval of 0.95 and a false discovery rate of 0.1 (Jacobsen et al. 2014). The second SNP characterization was carried out using the R package OutFLANK (Whitlock and Lotterhos 2015). The proportion of loci trimmed from both tails of the F_{ST} distribution was set to 5% (LeftTrimFraction = 0.05, RightTrimFraction = 0.05), the minimum heterozygosity required before including calculations from a locus was set to 0 (Hmin = 0) and the false discovery rate was 0.1 (qthreshold = 0.1). Only SNPs characterized as being under putative positive selection by both software packages were retained for downstream analyses.

Analysis of isolation by time requires the use of neutral markers that are under Hardy-Weinberg equilibrium (HWE) (Hendry and Day 2005). Therefore further filtering of the neutral SNP panel was performed in vcftools-gcc v.0.1.13 using the function --hwe 0.01.

To determine if the loci containing SNPs putatively under positive selection were contained in protein coding regions, sequences were BLASTed against the complete *Homarus americanus* transcriptome (McGrath et al. 2016) and a *J. edwardsii* transcriptome database (SRA Bioproject accession number: PRJNA386609) using BLAST+ v.2.2.29.

Queries with statistically significant e-values ($E < 10^{-6}$) and more than 84% identity were considered as valid alignments. Transcriptome sequences that provided significant alignments were annotated using the Trinotate pipeline (https://trinotate.github.io/) to determine if they aligned with any known protein domain.

Analyses of genetic diversity

Global F_{ST} values and confidence intervals for each SNP panel were estimated using the R package mmod v.1.3.2 (Winter 2012). Additionally, pairwise F_{ST} values between sampling sites and years as well as confidence intervals were calculated in the R package hierfstat v.0.04-22 (Goudet 2005). A false discovery rate correction (FDR) was applied to calculated p-values using the function "p.adjust" of the R package stats v.3.2.4 (R Core Team 2016). Additionally, a discriminant analysis of principal components (DAPC) was performed to determine the possible number of genetic clusters based on allele frequencies of all three sampling sites across three settlement months using the R package adegenet v.1.4-2 (Jombart 2008).

Analyses of Molecular Variance (AMOVA) (Excoffier et al. 1992) were performed in order to assess the amount of genetic variance explained by (1) settlement month, (2) sampling site and (3) the interaction of settlement month*settlement site using the R packages poppr v.2.1.1 (Kamvar et al. 2014) and ade4 v.1.7-4 (Dray and Dufour 2007).

Genetic diversity was quantified using the standardized individual heterozygosity (sh). This metric represents the proportion of heterozygous loci over the mean heterozygosity across all markers so that heterozygosity of all individuals are measured on the same scale (Coltman et al. 1999). The standardized individual heterozygosity was calculated for both SNP panels at (1) each site with all three settlement months combined and (2) at each settlement month at each site using the R package Rhh v.1.0.1 (Alho et al. 2010). Differences

between the maximum and minimum values of standardized individual heterozygosity were examined using a Mann-Whitney test. A Bonferroni correction was applied in order to account for multiple comparisons and α was set at the 0.05/4 level.

Kinship analysis

Kin relationships between pueruli pairs were inferred using the Loiselle's coancestry coefficient (Loiselle *et al*., 1995) implemented in the program GenoDive v2.0b27. The expected value of Loiselle's coefficient for full-sibs is 0.25 and 0.125 for half-sibs (Loiselle *et al*., 1995). Following Iacchei *et al*. (2013), we used the midpoints between those coancestry coefficients to generate bins as follows: full-sibs ($0.1875 < k \leq 0.375$), half sibs ($0.09375 < k \leq 0.1875$) and cousins ($0.047 < k \leq 0.09375$). Since the kinship analysis requires markers to be under Hardy-Weinberg equilibrium (HWE), we performed a further filtering of the neutral SNP panel using the function --hwe 0.01 of vcftools-gcc v0.1.13.

Isolation by time

IBT was estimated using Mantel tests between temporal and genetic distance using 9999 permutations within the R package ade4. The function "dist.gene" was used to compute genetic pairwise distance between samples using the R package ape v.3.4 (Paradis et al. 2004). The temporal distance matrix was calculated using the Euclidean distance of days between samples in the R package vegan.

Power analysis

We determined the power of the putatively neutral and outlier SNP markers for detecting genetic structure with the program POWSIM (Ryman and Palm 2006). Effective population size (N_e) was set to 3,000 individuals, according to the findings of Thomas & Bell (2013) and

considering 10 generations of drift (t). The generations of drift as well as N_e were used to calculate true F_{ST} values (Ryman and Palm 2006). Since the current version of POWSIM can only analyze 50 loci (Ryman and Palm 2006), only the statistical power of the first 50 neutral SNPs (out of 603) were calculated.

4.4 Results

Relationship between morphometric data and settlement time

The weight of individual puerulus differed significantly between sites and settlement months throughout the winter peak in settlement during 2012. However, the interaction between settlement site and month did not differ significantly (Table 4.2). Pueruli were lighter in weight in the most northerly site, Bicheno, and heavier at the mid sampling site, South Arm (Fig. 4.2). The post-hoc test identified significant differences between the weight of pueruli sampled from South Arm and the two other sampling sites only during August. Additionally, pueruli collected in Bicheno were significantly lighter in weight than those sampled from Recherche Bay and South Arm.

Table 4.2 Comparison of effects of settlement site, month and the interaction of both on pueruli weight-at-settlement. Significant differences are in **bold**

Variable	df	MS	F	*P*
Site	2	0.10975	57.753	**<2e-16**
Month	2	0.00593	3.122	**0.0477**
Site*Month	4	0.00257	1.352	0.2548
Error	118	0.00190		

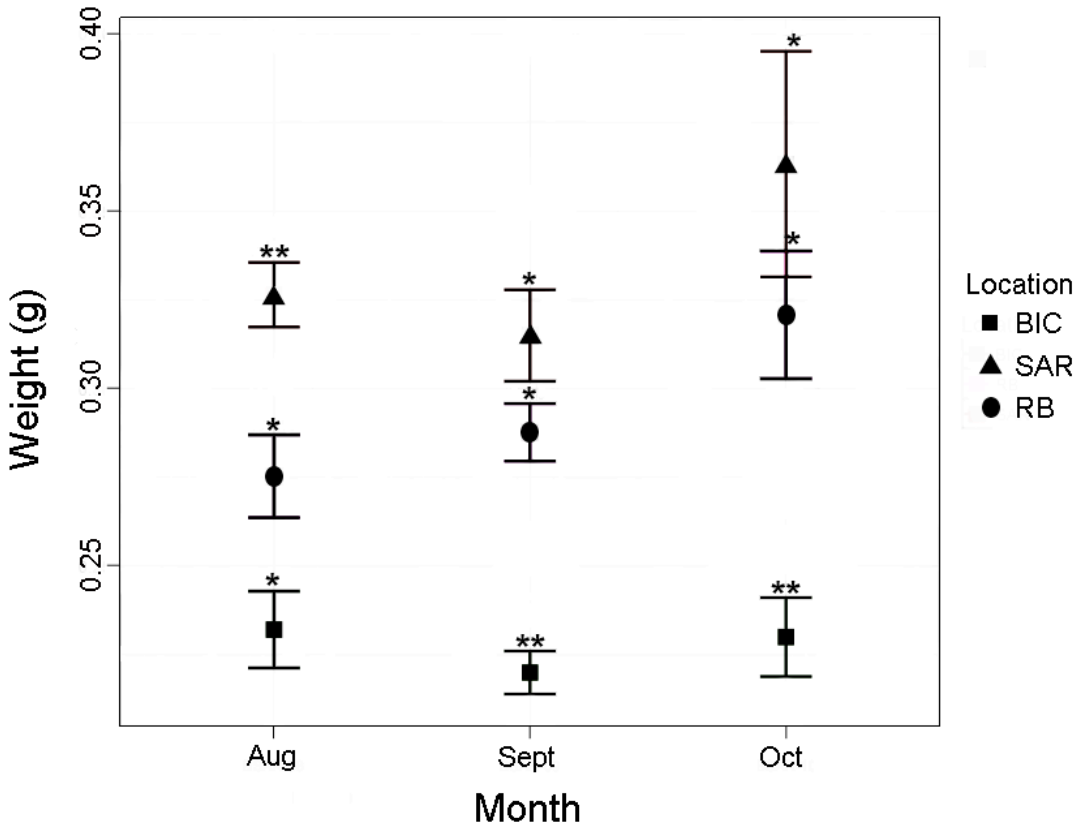


Fig. 4.2 Average individual weight (±SE) of pueruli settling into Bicheno (BIC), South Arm (SAR) and Recherche Bay (RB) during three months of study in 2012. No differences within sites were found. (*) Denotes non-significant differences between sites at each month; (**) denotes significant differences between sites at each month

Sequencing data and SNP characterization

Sequencing of four and a half Illumina HiSeq 2500 lanes yielded a total of 757.7 million reads, with an average of 3 million reads per sample. The reference catalog yielded 9,506 loci. After read alignment and variant calling, a total of 10,956 variable sites were obtained. After selecting one SNP per locus, filtering for MAF, LD and missing data the final number of SNPs was 1,194. The PCA plot showed each intralibrary technical replicate pair to be distributed close to one another. This indicates that sequencing samples across multiple lanes

did not introduce a large bias in the catalog building process (Fig. S1, Supporting information).

The SNP outlier identification performed in LOSITAN produced a panel of 85 SNPs under putative positive selection, 506 SNPs under putative balancing selection and 603 neutral SNPs. OutFLANK detected only two SNPs under putative positive selection, zero SNPs under putative balancing selection and 1,192 neutral SNPs. The final panel of 603 neutral SNPs and two SNPs under putative positive selection comprised loci identified by both software packages. HWE filtering of the neutral SNP panel detected an additional 190 SNPs not under HWE. Therefore the final panel of neutral markers was composed of 413 SNPs.

Analyses of genetic diversity

There was no evidence for genetic differentiation across all sites and settlement months in the neutral SNP panel (global F_{ST} = -0.001, n.s., Fig. S2, Supporting information). Pueruli settling into Recherche Bay during August were genetically distinct from individuals settling into Bicheno and South Arm (Table 4.3), corresponding to the phenotypic differences between sites described above. In addition, there were significant genetic differences between settlement months only at Recherche Bay, where individuals settling during August were genetically distinct from those settling during October (Table 4.3). Conversely, the SNPs under putative positive selection exhibited a stronger overall level of genetic differentiation (global F_{ST} = 0.0852, P = 0.0047). Pairwise F_{ST} values and AMOVA were not calculated for this SNP panel, since it was only comprised of two loci.

Differences between settlement months and the interaction of settlement month and site did not account for the overall variation in allele frequencies. Settlement site alone accounted for a small proportion of the variation between samples when analyzing the neutral SNP

panel (Table 4.4). This suggests that individuals settling at Recherche Bay (southeast Tasmania) at some sampling periods (August and October) could be sourced from different populations than those settling at Bicheno and South Arm (east Tasmania).

Table 4.3 Pairwise F_{ST} values between each site and month of settlement for the neutral SNP panel. Significant differences after False Discovery Rate (FDR) correction are in **bold**

	BIC Aug	BIC Sept	BIC Oct	SAR Aug	SAR Sept	SAR Oct	RB Aug	RB Sept
BIC Sept	0.0027							
BIC Oct	0.0021	0.0089						
SAR Aug	0.0029	0.0010	-0.0008					
SAR Sept	0.0021	0.0069	0.0019	0.0005				
SAR Oct	0.0027	0.0061	-0.0089	-0.0059	-0.0029			
RB Aug	0.0043	**0.0166**	**0.0160**	**0.0142**	0.0021	0.0187		
RB Sept	-0.0055	0.0117	-0.0014	-0.0019	0.0021	0.0046	0.0028	
RB Oct	0.0009	0.0047	0.0026	0.0014	0.0055	-0.0081	**0.0205**	0.0015

Table 4.4 Analysis of molecular variance (AMOVA) results using genetic distance as a function of settlement site. Significant differences are in **bold**

Variance component	Variance	% total	*P*	Φ
Neutral SNP panel				
Variations between settlement month	-0.006	-0.016	0.572	$-1.60e^{-04}$
Variations between settlement site	0.063	0.158	**0.034**	0.002
Variations between month within site	0.002	0.004	0.516	$1.59e^{-03}$

The overall genetic diversity of the neutral SNP panel exhibited no difference across sampling sites or months (Fig. 4.3).

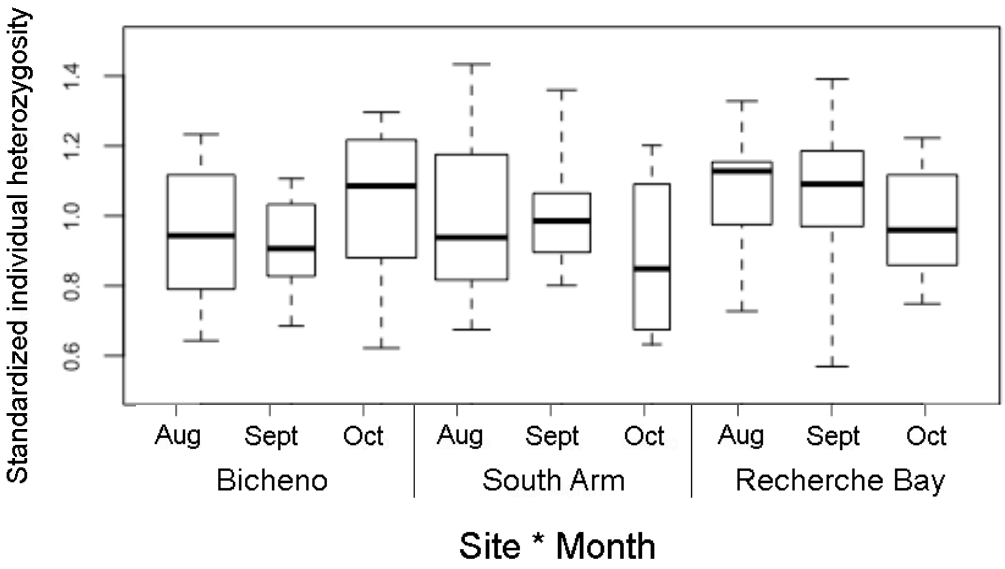


Fig. 4.3 Median standardized individual heterozygosity (sh) at each site during each settlement month when analyzing the neutral SNP panel. Whiskers represent minimum and maximum values

Finally, transcriptome sequences from the eyestalk, optical nerve, green gland and hepatopancreas of *J. edwardsii* exhibited significant hits with one locus containing SNPs under putative positive selection. However, the transcriptome sequences could not be successfully annotated to any gene. Additionally, no significant hits between the panel of SNPs under putative positive selection and the *Homarus americanus* transcriptome were found.

Kinship analysis

A total of 5,356 pairwise comparisons were computed, out of which 197 individuals were found to be third degree relatives or cousins. No spatial or temporal patterns of kinship were detected. Also, no full or half-sibs were identified in the dataset.

Power analysis

The overall power estimate of 50 neutral SNPs from the 1,000 simulated runs was 0.4210 for Chi-square and 0.3980 for the Fisher approach, based at a significance level of $\alpha = 0.05$.

Evidence for chaotic genetic patchiness and isolation by time

In support of the small F_{ST} differentiation between settlement months in Bicheno and South Arm, the isolation by time analyses showed no significant correlation between genetic distance and distance between recruitment time (Table 4.5). Similarly, the significant F_{ST} values found among the Recherche Bay sampling periods were confirmed by the significant Mantel test at this site. Pueruli settling into Recherche Bay during the same month were genetically more similar to each other than individuals settling one or two months apart, as shown by the slight but significant regression (Fig. 4.4). This suggests the existence of slight chaotic genetic patchiness at Recherche Bay.

Table 4.5 Results of the Mantel tests between temporal and genetic distance at each settlement site using the neutral SNP panel

Settlement site	r	*P*
Bicheno	-0.004	0.5418
South Arm	-0.153	0.9801
Recherche Bay	0.087	**0.0346**

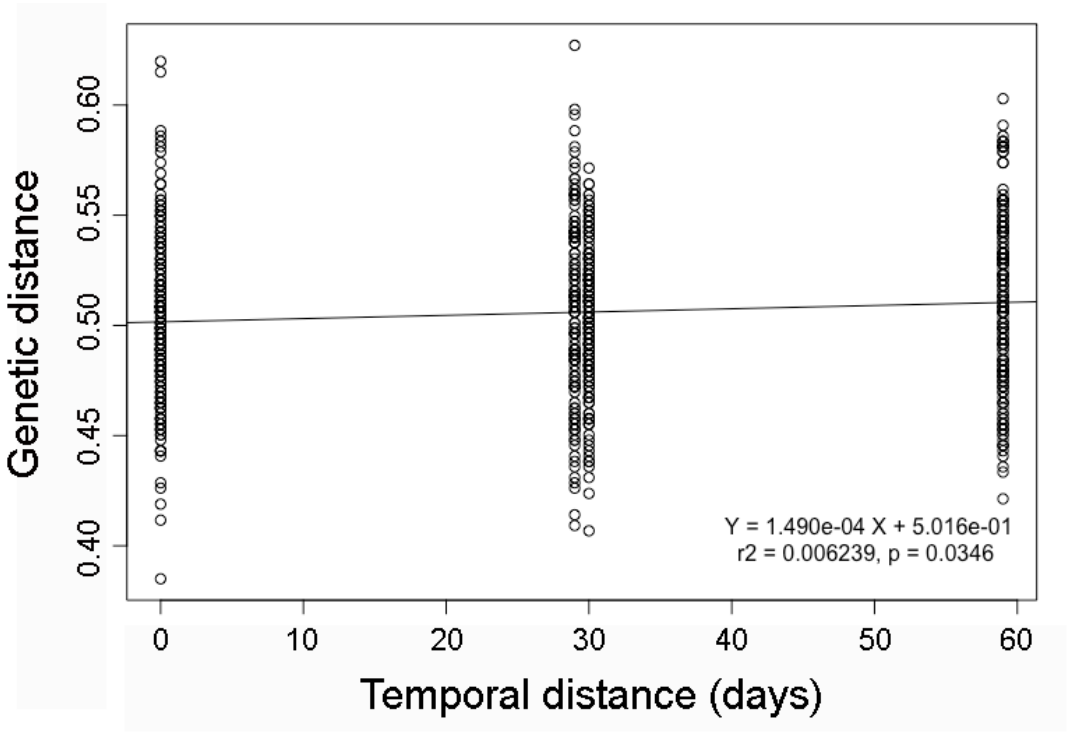


Fig. 4.4 Relationship between genetic distance at the neutral SNP panel and temporal distance at Recherche Bay. On the x-axis, zero corresponds to samples collected within the same month, 29 corresponds to genetic distance between individuals settling during August and September, 31 corresponds to genetic distance between individuals settling during September and October, 60 corresponds to genetic distance of individuals settling between August and October

4.5 Discussion

The present study examined puerulus size and genetic signatures during the winter settlement peak at three sites on the east coast of Tasmania, Australia. Significant differences in weight between settling pueruli were evident amongst the study sites. Pueruli settling into the northernmost site, Bicheno, were consistently smaller than those arriving to the other two sampling sites. Despite the overall low levels of genetic differentiation at neutral markers there was some evidence for chaotic genetic patchiness at the southernmost site. Finally, genetic differentiation found at SNPs putatively under selection could indicate that the observed phenotypic differences across sites and months were driven by selection in response to environmental factors.

Dispersal history driving phenotypic plasticity

The observed phenotypic differences between pueruli settling into three sites in Tasmania during the study period could be caused by maternal influences into offspring quality, the dispersal history of settling individuals, environmental differences at the sites, or differences in the food and environmental conditions experienced through larval development. The former case would imply that larger females could have sourced settlers into South Arm during the three study months. However, given the little genetic differentiation found herein, this hypothesis is unlikely.

Variations in size-at-settlement, measured as puerulus carapace length, have been observed in *J. edwardsii* arriving during different years and seasons into the same site (Booth 1979; Booth and Tarring 1986). For example, stage 1 pueruli settling during winter on the northwest of New Zealand were significantly bigger than those settling during other times of the year (Booth and Tarring 1986). Also, in a three-year study, Booth (1979) found significant differences in size-at-settlement of *J. edwardsii* arriving into one site on northeast

New Zealand between most years and months of settlement. The authors hypothesized that the phenotypic differences were caused by different cohorts of phyllosoma that metamorphosed into pueruli and settled simultaneously and to different parental stocks (Booth 1979).

The lack of neutral genetic structure found in the present study indicates that it is unlikely that the phenotypic variation is directly linked to different larval sources. This leads us to hypothesize that the observed morphological differences are due to phenotypic plasticity possibly driven by dispersal history. The environment has profound effects on the metabolism and growth of larval lobsters and this influence during early life can be carried over into subsequent stages (Green et al. 2014). *J. edwardsii* settling into east Tasmania can spend up to two years as phyllosoma larvae, encountering a large gradient of environmental factors (Bruce et al. 2007). Temperature and food availability are the main factors affecting growth rates of larvae and can determine settlement success (Pechenik 2006). Therefore, individuals that encountered temperature and food conditions that promoted growth during the late phyllosoma stage could have higher accumulation of energy reserves and metamorphose into larger pueruli (Jeffs et al. 2001).

If the significant differences in pueruli weight across sites were caused by differences in temperature and food during development, the persistence of size differences to sites requires that larvae were travelling in cohorts and experiencing different temperature and nutrition regime, for at least the last part of their larval life. This would lead to a variation in developmental conditions of individuals that settle across a latitudinal gradient. Collective dispersal is a possible cause of chaotic genetic patchiness (Eldon et al. 2016) and has been previously proposed to occur in spiny lobsters (Iacchei et al. 2013; Funes-Rodríguez et al. 2015). Genetic patchiness and high levels of within-site kinship in individuals of the California spiny lobster, *Panulirus interruptus*, suggested the potential for larvae spending

the pelagic phase as a group and settling together (Iacchei et al. 2013). This was later supported by a study assessing the seasonal distribution of *P. interruptus* phyllosoma in relation to hydrographic conditions (Funes-Rodríguez et al. 2015).

While larval cohesiveness could be a plausible method of dispersal in *J. edwardsii*, individuals of different ages could also congregate near settlement grounds, spending only the last period of their PLD together. Late-stage *J. edwardsii* phyllosoma have been observed to congregate beyond the shelf break in oceanic waters, possibly through a combination of oceanic advection and horizontal swimming (Chiswell and Booth 1999). Differences in age-at-settlement and dispersal history could possibly result in differences in size-at-settlement in *J. edwardsii* found in the present study. A mixture of collective dispersal and convergence of larvae at later stages has been described in a reef fish cohort settling into the same site (Shima and Swearer 2016). If this also occurs in *J. edwardsii*, then the genetic signature of settling individuals will be more randomly distributed, with no genetic patchiness, as seen in Bicheno and South Arm, but with a large variation in phenotypes of settlers due to divergent dispersal histories.

Finally, the temporal differences in pueruli weight could be attributed to differences in environmental conditions related to time of hatching. A latitudinal gradient of incubation period due to differences in temperature has been described in the Dungeness crab, *Cancer magister* (Shirley et al. 1987) where individuals hatching at lower temperatures were larger than those hatching in warmer waters. Also, if time and place of hatching can determine where larvae will be advected, this would result in different larval cohorts experiencing divergent environmental conditions. It is therefore possible that females spawning at different times of a reproductive season can produce offspring with slightly different phenotypes (Kunisch and Anger 1984; Shirley et al. 1987). This could have been the case in the present study, where individuals recruiting during different months exhibited significant differences

in weight at settlement. The same trend was observed in the green crab, *Carcinus maenas*, where larval size at metamorphosis and survival of the first crab instar under starvation varied significantly across four supply events (Rey et al. 2016). Individuals sampled later in the season exhibited larger size than those sampled earlier, probably due to more favorable environmental conditions encountered by the former group (Rey et al. 2016).

The question that arises here is whether fitness in *J. edwardsii* settlers is related to size-at-settlement, leading to successful recruitment. Although there is a general belief is that larger offspring are fitter than smaller offspring, the link between size and fitness is not completely understood (Marshall et al. 2006; Gimenez 2006). Larger individuals generally exhibit improved anti-predatory behavior, are bolder, can forage for a longer time and better resist periods of starvation (Dingeldein and White 2016; Johnson et al. 2017). However it is often the environment (biotic and abiotic factors) that ultimately provides the selective pressure and determines survival (Marshall et al. 2006).

If smaller size represents a disadvantage for post-settlement survival, it would then result into fewer pueruli recruiting to the fishery and a lower productivity from Bicheno compared to the other sampling areas studied here. This is actually the case. Average counts of puerulus at Bicheno are substantially higher than the other sites (unpublished data), yet the Bicheno region has the lowest fishery productivity in Tasmania (Gardner et al. 2015a). Therefore, it is possible that in *J. edwardsii*, smaller pueruli are less fit and less likely to recruit into the fishery.

Evidence for chaotic genetic patchiness and natural selection

Lack of neutral genetic differentiation between sampling events suggests that populations sourcing pueruli into the east coast of Tasmania during winter months are genetically homogenous, however there is a slight indication of genetic differentiation at the

southernmost site. Lack of overall genetic divergence but significant differences at smaller spatial or temporal scale is typical of chaotic genetic patchiness (Eldon et al. 2016). This indicates that the patchy genetic structure between South Australia and Tasmania evident from chapter 3 of this thesis can also be detected at this fine geographic and temporal scale.

The significant genetic differentiation between pueruli settling during August and October at Recherche Bay could indicate the existence of different populations (or groups of individuals) sourcing pueruli during the winter months. The slight differences in individual weights between months found and that size-at-settlement could be a signature of each of these populations. This finding is in line with predictions of the IBT hypothesis, which proposes that the existence of early and late spawners within a population drives a patchy genetic structure of the settlers (Hendry and Day 2005).

The seasonal reproduction and the latitudinal variation in egg-hatching period of *J. edwardsii* are traits that could make the species prone to genetic patchiness due to isolation by time. Since phyllosoma larvae move to the water surface shortly after hatching, the prevailing wind conditions can transport a whole larval assemblage offshore (Booth and Phillips 1994). Collective larval transport has been described in *P. interruptus* phyllosoma, where individuals released during each spawning event use the same transport mechanisms and thus maintain cohesiveness during dispersal and settle together (Funes-Rodríguez et al. 2015). In the case of *J. edwardsii*, if different groups of ovigerous females aggregate at different days to spawn (McKoy and Leachman 1982) it could result in a patchy distribution of larvae in the ocean (Rimmer 1980) that could metamorphose and settle together.

Temporal genetic heterogeneity in recently settled pueruli has also been reported in a closely related lobster species. Allozyme analysis of monthly collections of recently settled western rock lobster pueruli showed differences between individuals settling early and late in the season (Johnson and Wernham 1999). However, the pattern was only observed at one of

three settlement years, suggesting the possibility of sweepstakes reproductive success, SRS, (Hedgecock 1994) or natural selection (Kennington et al. 2013c) rather than IBT. Given that the genetic diversity through time did not vary significantly in the present study, we conclude that SRS is not a possible cause of the observed pattern of genetic differences between months. The hypothesis of SRS driving the temporal genetic structure in settling pueruli was also rejected in chapter 3 of this thesis based on the absence of significant differences in the genetic diversity between years and lack of genetically distinct cohorts of pueruli arriving into this site within a year. Therefore, in the present study IBT and natural selection could be two possible mechanisms behind the genetic and phenotypic divergence of settling pueruli.

The differences in allele frequencies at the panel of SNPs under putative positive selection and the fact that the phenotypic differences in recruits between sites persist across months imply that natural selection could be driving the variation in pueruli phenotypes at each site. It is known that allelic variation at specific loci can control gene expression and produce phenotypic differences in the presence of divergent environments (Beldade et al. 2011). The three sampling sites studied here vary slightly in environmental conditions: there is a latitudinal gradient in water temperature, being warmer at Bicheno than the two other sampling sites. Also, South Arm is located at the mouth of the Derwent River and experiences lower salinity than the other sampling sites. Therefore the selective pressure on recruits could also differ according to location. Yet, variable post-settlement selection across time and space on a single pool of larvae could also drive the observed phenotypic variation. This is in line with the findings of chapter 3 of this thesis, which provided evidence of post-settlement selection in Bicheno varying in strength across years.

4.6 Conclusions

The observed variation in size-at-settlement of pueruli observed in the present study could be due to phenotypic plasticity with differences driven by development under different environmental conditions through its protracted PLD. The hypothesis of cohorts of larval *J. edwardsii* maintaining cohesiveness and experiencing the same environmental conditions is a likely explanation for the observed phenotypic differences between settlement months and sites in the lack of overall neutral population differentiation. Whether larvae arriving into different settlement sites originated from different regions is a question that remains to be answered. Given the differences in size-at-settlement described in the present study it would be interesting to investigate if the patterns are consistent over a larger area, across southern Australia and also to determine their implications for puerulus fitness and recruitment success.

The lack of neutral population differentiation suggests that neither the time nor place of reproduction acts as a biological barrier to dispersal of *J. edwardsii* populations on the east coast of Tasmania. However differences in regions of the genome under putative positive selection across the spatial and temporal scales also confirm that this species undergoes weak pre-settlement natural selection, and demonstrate that this occurs at a fine geographic and temporal scale.

4.7 Acknowledgements

Funding for this research was provided by an Australian Research Council Linkage Project grant (Project No. LP120200164) and Australian Research Council Discovery Project grant (Project No. DP150101491). We thank Ivan Hinojosa, Ruari Colquhoun, David Fallon, Kylie Cahill, Graeme Ewing and Sarah Pyke for puerulus collections in Tasmania. Thanks to

Brooke Whitelaw and Daniel Laird for laboratory assistance. Special thanks to Karen J. Miller who contributed to the original project idea.

4.8 Supplementary information

Table S1 Differences between mean individual weight per month and site of settlement and confidence intervals at a 95% confidence level based on the "Tukey's Honest Significant Difference" method. Values in bold are significant after adjustment for multiple comparisons

Site (Settlement month)	Mean difference	Lower limit	Upper limit	*P*-adj
RB (Aug) - BIC (Aug)	0.0432	-0.0044	0.0908	0.1066
SAR (Aug) - BIC (Aug)	0.0944	0.0503	0.1385	**0.0000**
BIC (Sept) - BIC (Aug)	-0.0121	-0.0581	0.0339	0.9957
RB (Sept) - BIC (Aug)	0.0556	0.0060	0.1052	**0.0160**
SAR (Sept) - BIC (Aug)	0.0829	0.0321	0.1337	**0.0000**
BIC (Oct) - BIC (Aug)	-0.0021	-0.0497	0.0455	1.0000
RB (Oct) - BIC (Aug)	0.0887	0.0391	0.1382	**0.0000**
SAR (Oct) - BIC (Aug)	0.1312	0.0456	0.2168	**0.0001**
SAR (Aug) - RB (Aug)	0.0512	0.0041	0.0982	**0.0223**
BIC (Sept) - RB (Aug)	-0.0553	-0.1041	-0.0065	**0.0141**
RB (Sept) - RB (Aug)	0.0124	-0.0398	0.0646	0.9979
SAR (Sept) - RB (Aug)	0.0397	-0.0137	0.0930	0.3219
BIC (Oct) - RB (Aug)	-0.0453	-0.0956	0.0050	0.1127
RB (Oct) - RB (Aug)	0.0454	-0.0068	0.0976	0.1419
SAR (Oct) - RB (Aug)	0.0880	0.0009	0.1751	0.0458
BIC (Sept) - SAR (Aug)	-0.1065	-0.1519	-0.0611	**0.0000**
RB (Sept) - SAR (Aug)	-0.0388	-0.0879	0.0103	0.2433
SAR(Sept) - SAR (Aug)	-0.0115	-0.0618	0.0388	0.9984
BIC (Oct) - SAR (Aug)	-0.0965	-0.1436	-0.0494	**0.0000**
RB (Oct) - SAR (Aug)	-0.0057	-0.0548	0.0433	1.0000
SAR (Oct)-SAR (Aug)	0.0368	-0.0485	0.1221	0.9086
RB (Sept) - BIC (Sept)	0.0677	0.0169	0.1184	**0.0016**
SAR (Sept) - BIC (Sept)	0.0950	0.0431	0.1469	**0.0000**
BIC (Oct) - BIC (Sept)	0.0100	-0.0388	0.0588	0.9993
RB (Oct) - BIC (Sept)	0.1008	0.0500	0.1515	**0.0000**
SAR (Oct) - BIC (Sept)	0.1433	0.0571	0.2296	**0.0000**
SAR (Sept) - RB (Sept)	0.0273	-0.0278	0.0825	0.8216
BIC (Oct) - RB (Sept)	-0.0577	-0.1099	-0.0055	**0.0188**
RB (Oct) - RB (Sept)	0.0331	-0.0210	0.0871	0.5915

Table S1 continued

Site (Settlement month)	Mean difference	Lower limit	Upper limit	*P*-adj
SAR (Oct) - RB (Sept)	0.0756	-0.0126	0.1639	0.1560
BIC (Oct) - SAR (Sept)	-0.0850	-0.1384	-0.0316	**0.0001**
RB (Oct) - SAR (Sept)	0.0058	-0.0494	0.0609	1.0000
SAR (Oct) - SAR (Sept)	0.0483	-0.0406	0.1373	0.7345
RB (Oct) - BIC (Oct)	0.0908	0.0386	0.1430	**0.0000**
SAR (Oct) - BIC (Oct)	0.1333	0.0462	0.2205	**0.0001**
SAR (Oct) - RB (Oct)	0.0426	-0.0457	0.1308	0.8419

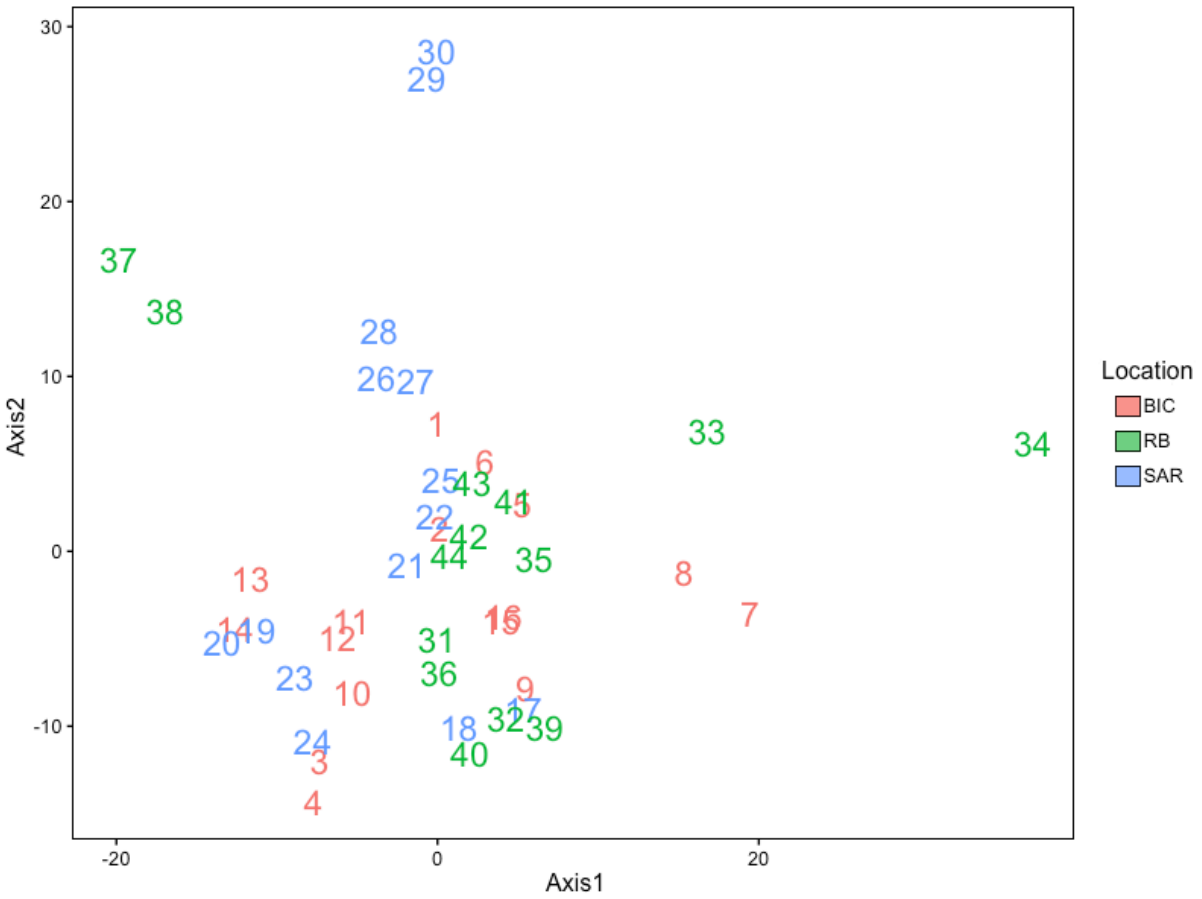


Fig. S1 Principal component analysis based on allele frequencies of all 22 replicate pairs from Bicheno (red numbers), South Arm (blue numbers) and Recherche Bay (green numbers). Consecutive numbers belong to each technical replicate pair

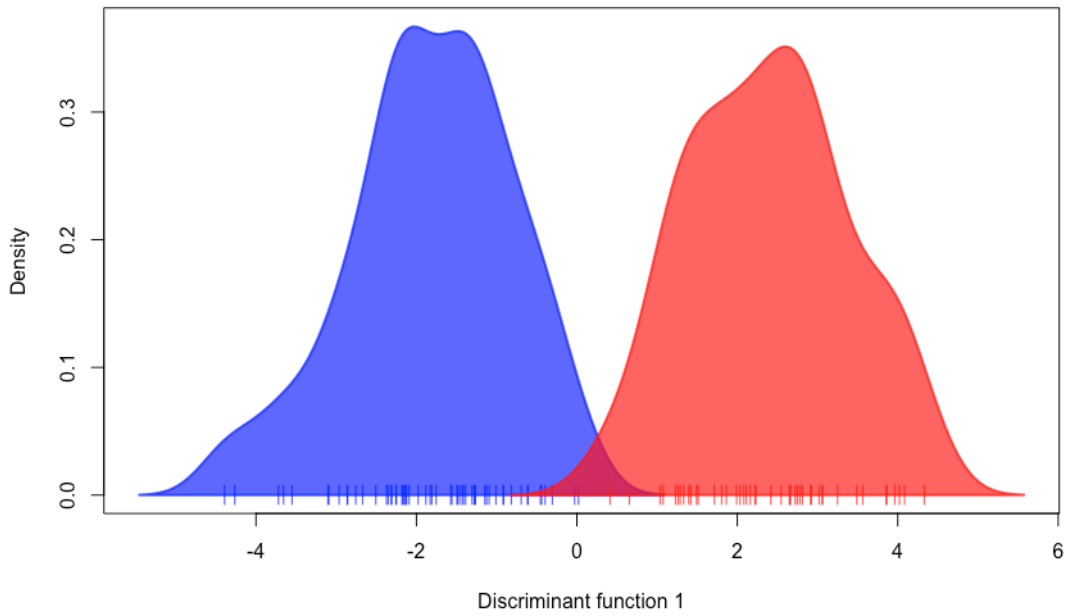


Fig. S2 First principal component resulting from the DAPC using pueruli settling into Bicheno, South Arm and Recherche Bay during three consecutive winter months during 2012 with the neutral SNP panel

Chapter 5: Preliminary assignment of *Jasus edwardsii* post-larvae to population of origin

5.1 Abstract

Knowledge of source locations of newly settled recruits can help aid fisheries management decisions directed towards protecting vulnerable populations. The Southern rock lobster, *Jasus edwardsii*, is considered a panmictic population within Australia. Despite this, the fisheries management is carried out independently in the three states where this resource is targeted commercially. Based on predictions of larval transport simulations, we hypothesized that South Australian populations are an important source of recruits for eastern areas, including coastal areas within the states of Tasmania and Victoria. The present study aimed to conduct population assignment of pueruli into adult sampling sites along southeast Australia. Adult *J. edwardsii* were collected from three sampling areas that have been suggested as source locations of newly settled recruits: the northern fishing zone of South Australia, the southern fishing zone of South Australia and southeast Tasmania. Additionally, pueruli samples from one sampling site in South Australia and three in Tasmania were collected. Given that there is interannual variability in settlement, pueruli from two consecutive sampling years, 2012 and 2013, were included in the analyses. All three adult sampling sites were assigned as putative sources of recruits for all sampling sites, except Recherche Bay, in both years. This is in line with predictions of the larval transport simulations, however caution should be taken when interpreting the results due to potential bias and lack of power introduced by not sampling all putative source populations.

5.2 Introduction

Knowledge of dispersal patterns as well as migration rates in populations of marine organisms of commercial concern can help inform managerial decisions (Hastings and Botsford 2006). It can lead to delineation of management units based on the biology of the species, rather than on geopolitical boundaries (Kerr et al. 2014; Nicolle et al. 2016). Disregarding a population's genetic structure can result in inaccurate estimates of spawning stock biomass, annual yield and annual recruitment (Reich and DeAlteris 2009). Empirical and simulated data show that unequal larval dispersal make some sub-populations along a species distribution more vulnerable to overfishing (Fu and Fanning 2004; Truelove et al. 2015). Sub-populations where the net export of individuals is larger than the net import of individuals, commonly known as "source", should be managed differently from "sink" populations, where the net import of individuals is greater (Chiswell and Booth 2008).

The Southern rock lobster, *Jasus edwardsii*, exhibits the longest pelagic larval duration (PLD) amongst all the species within the palinurid family. Its larval phase extends between 12 to 24 months (Booth 1994) and flat-bodied larvae are well adapted for passive drifting. Larval dispersal simulations of *J. edwardsii* in Australia suggest that both self-recruitment within a region and migration of larvae between regions are possible (Bruce et al. 2007). There is a general eastward flow of currents displacing larvae and promoting genetic exchange across the species distribution. Additionally, mesoscale eddies can entrain larvae and promote self-recruitment, particularly in the southwest of Western Australia (Bruce et al. 2007). The southern zone of South Australia comprises a highly productive region and due to the oceanographic features, namely the South Australian Current (SAC), the Zeehan Current (ZC) (Fig. 1.1) and mesoscale eddies, it is thought to provide recruits into both areas north and south of this region. Westward movement of larvae from Tasmania has also been hypothesized based on temporal oceanic features. There is an oceanic cyclone occurring on

average every 18 months off the west coast of Tasmania, which could potentially transport phyllosoma from the western coast of Tasmania northwards to the states of South Australia and Victoria (George 1997). However larval transport simulations have hypothesized low levels of recruitment into South Australia from east and west Tasmania during some years (Bruce et al. 2007).

Numerous studies assessing population structure of *J. edwardsii* using a range of genetic tools have failed to detect genetic differentiation within Australia (Ovenden et al. 1992; Morgan et al. 2013; Villacorta-Rath et al. 2016) and the population is considered as a single stock. Despite being a panmictic species within Australia, the fishery is managed separately in three states where this resource is targeted commercially (South Australia, Victoria and Tasmania). Annual recruitment is highly variable due to environmental factors affecting larvae prior to settlement (Linnane et al. 2010b) and possibly due to variability in reproductive success of the spawning stock (Booth 1994). This uncertainty hinders predictions of the effects of harvesting strategies on the population levels, since models are highly dependent on good estimates of annual recruitment (Gardner et al. 2015b). The effect of recruitment variation on economic yield is substantial. For example, the difference between high and low recruitment scenarios have an economic impact of approximately AUD$100 million (McGarvey et al. 2016). This highlights the importance of understanding the combination of biological and environmental factors behind yearly fluctuations in recruitment, including investigating potential sources of recruits.

Genomic tools have been recognized as a powerful approach to advance the understanding of the ecology of commercial marine species. This knowledge can be used to help advise fisheries managers (Bernatchez et al. 2017). Among them, assignment tests are increasingly used in commercial species since they have direct applications for food traceability (Larraín et al. 2014; Ogden and Linacre 2015; Villacorta-Rath et al. 2016), to

protect consumers from seafood that constitute a threat to public health (Bernatchez et al. 2017) and to detect illegal fishing (Nielsen et al. 2012a). Assignment tests determine an individual's most likely population of origin based on allele frequencies (Rannala and Mountain 1997). In the present study, based on the simulated larval transport models (Bruce et al. 2007), we tested the hypotheses that South Australian adult populations account for a large proportion of recently settled recruits in Tasmania and that there is interannual variation in the source of recruits. The main objective of this study was to assign pueruli settling into South Australia and Tasmania during two consecutive years to their most likely population of origin based on allele frequency data using double digest restriction-site associated DNA sequencing (ddRADseq).

5.3 Methods

Sample collection from putative source and sink populations

Pleopod clips of adult lobsters from five putative source populations in South Australia and Tasmania were collected using baited lobster pots during 2013 and 2015 and immediately preserved in 90% ethanol. All female specimens collected in South Australia and Tasmania had reached sexual maturity, as they were either berried or possessed long setae. Male specimens from South Australia were above the legal size, while individuals from Tasmania were either legal-sized (110 mm) or up to 10 mm smaller. Adult specimens were collected from five sampling sites: the southern fishing zone of South Australia, the northern fishing zone of South Australia, Betsey Island (Tasmania), Bruny Island (Tasmania), Taroona (Tasmania) (Fig. 5.1, open circles). Given that the three sampling sites in Tasmania were located approximately 20 km from each other and no significant genetic differentiation was detected between them (Table S1, Supporting information) they were pooled into one site, hereafter referred to as "southeast Tasmania".

Additionally, pueruli from four putative sink populations in South Australia and Tasmania were sampled in puerulus collectors deployed permanently at four sites: Cape Jaffa (South Australia), Recherche Bay (Tasmania), South Arm (Tasmania) and Bicheno (Tasmania) (Fig. 5.1, filled circles). After collection, whole pueruli were preserved in 90% ethanol. Since chapter 2 of this thesis provided evidence for variability in genetic structure and diversity of the settling cohorts among years, samples from two consecutive years (2012 and 2013) from all pueruli sampling sites except Recherche Bay were considered separately in the analysis (Table 5.1).

DNA extraction and ddRADseq library preparation

DNA was extracted from tissue from pleopod clips, in the case of adult lobster specimens, and the horns and legs of each individual, in the case of pueruli specimens, using a DNeasy Blood and Tissue kit (Qiagen) following the manufacturer's instructions. Subsequently, DNA concentration was quantified using a Qubit® 2.0 Fluorometer (Life Technologies). DNA integrity was determined through gel electrophoresis to verify high molecular weight DNA in all samples (>1,000 base pairs [bp] or higher).

ddRADseq libraries were prepared following the protocol described in Villacorta-Rath et al. (2016) using a gel size selection of 400-600 bp in order to maximize the overlapping region among libraries. A total of three interlibrary technical replicates and seven intralibrary technical replicates were distributed across the ddRADseq libraries to test for batch effects of sequencing across different lanes (Mastretta-Yanes et al. 2015). Electrophoretic assays were performed on all ddRADseq libraries by AGRF using an Agilent BioAnalyser (Agilent Technologies) to determine their exact molecular weight range and library concentration. A total of 291 samples were sequenced across five lanes at the Australian Genome Research Facility (AGRF) on the Illumina HiSeq 2500 platform using a 100 bp single end kit.

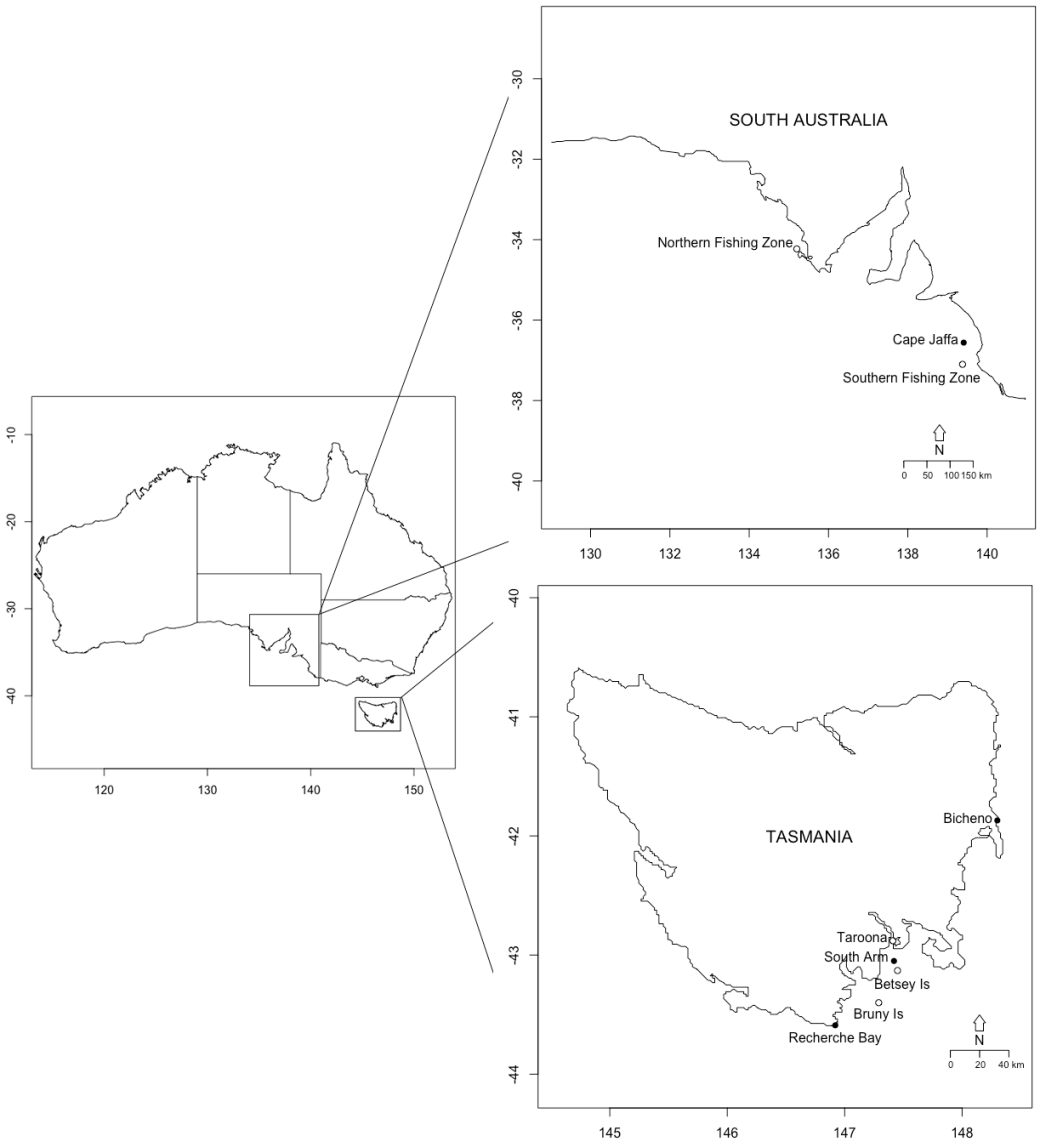


Fig. 5.1 Sampling sites in South Australia and Tasmania. Filled circles correspond to puerulus collectors sites. White circles correspond to adult sampling sites

Table 5.1 Sampling sites for adult and puerulus *J. edwardsii* collected in South Australia and Tasmania. Sampling year, number of individuals sequenced (N_i), final number of individuals included in analyses (N_f) after filtering for missing data and removal of replicates for each sampling site. Observed heterozygosity (H_O) and inbreeding coefficient (F_{IS}) are reported for each site

State	Sampling site	Stage	Sampling year	N_i	N_f	H_O	F_{IS}
South Australia	Northern fishing zone	Adult	2015	27	20	0.546	-0.480
	Southern fishing zone	Adult	2015	30	16	0.540	-0.459
	Cape Jaffa	Puerulus	2012	18	18	0.553	-0.489
			2013	9	9	0.576	-0.505
Tasmania	Southeast Tasmania	Adult	2013	64	55	0.554	-0.495
	Recherche Bay	Puerulus	2012	33	30	0.559	-0.490
	South Arm	Puerulus	2012	33	25	0.513	-0.425
			2013	27	27	0.510	-0.423
	Bicheno	Puerulus	2012	30	30	0.549	-0.483
			2013	20	20	0.562	-0.513

Analyses of raw sequencing data and reference catalogue building

An initial quality check of raw indexed data was performed using FastQC v.0.10.1. Data was then demultiplexed using the "process_radtags" protocol from Stacks v.1.29 (Catchen et al. 2011) and hard trimmed to 75 bp to ensure that Phred Quality Score (Q Score) of all reads was above 30. Demultiplexed libraries were filtered for bacterial and viral content using Kraken-gcc v.0.10.4 (Wood and Salzberg 2014). Filtered reads were processed through the rad-loci pipeline described in Villacorta-Rath et al. (2016) to obtain a catalogue of reference loci (https://github.com/molecularbiodiversity/rad-loci).

Since the total number of individuals sampled in this study was 291, clusters with a depth of 270 or more reads were retained initially in order to ensure that most samples would be represented with at least one read. Within each cluster a 4% mismatch (3 bp) between reads was allowed, which means that reads shared 96% identity. Subsequently, clusters that had less than two and more than 16 members (or alleles) were discarded. A second round of clustering of the remaining clusters was performed at a 96% identity and a subsequent filtering of clusters that were not composed of a minimum of two and maximum of 16 members. Finally, samples were mapped back to the identified loci allowing for 30% missing data.

Read alignment, variant calling and SNP filtering

Individual filtered reads of all replicates were aligned to the reference catalog using bwa-intel v.0.7.12 (Langmead and Salzberg 2012). Subsequently variant calling of aligned reads was performed through the Genome Analysis Toolkit (GATK) v.3.3_0 (McKenna et al. 2010). A further correction of the variant call format (vcf) file obtained from GATK was performed to ensure the accuracy of the reference and alternate allele calls and to filter out false positives. In the absence of a reference genome, the correction was made by calculating

the ratio between the highest quality score over depth ("QD" in vcf file) and lowest QD. If one sample at a specific position had a ratio threshold of 10, which corresponds to a 10% error on a Phred scale, it was substituted by missing data for that SNP.

SNP filtering was performed in vcftools-gcc v.0.1.13 to ensure that only bi-allelic data was present (--min-alleles 2, --max-alleles 2), to remove SNPs that were potentially in linkage disequilibrium (--min-r2 0.2), to discard SNPs with a minor allele frequency (MAF) of less than 10% (--maf 0.1) and to ensure that the minimum SNP depth was 5 (--minDP 5). The MAF threshold of 10% was chosen because the downstream analyses of population assignment are model-based approaches, which are highly dependent on frequencies of rare and common alleles (Linck and Battey 2017). Setting a more stringent MAF cut-off can help reveal population structure that can be masked with a lower MAF (Linck and Battey 2017). The maximum amount of missing data for each locus was set to 15% (--max-missing 0.85) and individuals with more than 20% missing data were removed from subsequent analyses. Finally, only one SNP per locus (--thin 75) was retained.

A total of 14 intra and interlibrary technical replicates remained after removal of individuals with more than the missing data threshold. A principal components analysis (PCA) was performed to visualize the spatial distribution of replicated samples.

In order to minimize false positives in neutral and outlier loci identification, two software packages were used to characterize SNPs. The first SNP characterization was performed in LOSITAN (Antao et al. 2008) using 100,000 simulations, a confidence interval of 0.95, and a false discovery rate of 0.1 (Jacobsen et al. 2014). LOSITAN uses an F_{ST}-outlier approach that identifies loci as outlier candidates when they exhibit too high or too low F_{ST} compared to neutral expectations (Antao et al. 2008). The second SNP characterization was carried out using the R package OutFLANK (Whitlock and Lotterhos 2015). The proportion of loci trimmed from both tails of the F_{ST} distribution was set to 5% (LeftTrimFraction = 0.05,

RightTrimFraction = 0.05), the minimum heterozygosity required before including calculations from a locus was set to 0 (Hmin = 0) and the false discovery rate was 0.1 (qthreshold = 0.1). Only SNPs characterized as being neutral, under putative positive selection and under putative balancing selection by both software packages were retained for downstream analyses.

Analyses of genetic diversity and differentiation

The level of observed (H_O) and expected heterozygosity (H_E) in each sampling site was calculated using the R package adegenet v.2.0.1 (Jombart 2008). Global F_{ST} and confidence intervals for the neutral SNP panel were estimated using the R package mmod v.1.3.2 (Winter 2012). Additionally, pairwise F_{ST} values between sampling sites as well as confidence intervals were calculated in the R package hierfstat v.0.04-22 (Goudet 2005). A false discovery rate correction (FDR) was applied to calculated p-values using the function "p.adjust" incorporated within the R package stats v.3.2.4 (R Core Team 2016).

Assignment tests

Assignment of groups of individuals into putative population of origin was performed using the software GeneClass2 v.2.0 (Piry 2004) with the Bayesian assignment method developed by Rannala and Mountain (1997). This test allows detection of immigrants up to two generations in the past, even when the overall genetic differentiation between populations is low (Rannala and Mountain 1997). A group of individuals (or sampling site) was considered immigrant when the probability of being assigned to a population was $P < 0.05$. Two different datasets were analyzed, one where the "individuals to-be-assigned" were pueruli collected during the year 2012 and the other where the "individuals to-be-assigned"

were pueruli collected during the year 2013. In both cases the reference populations were the adult *J. edwardsii* collected from the three sampling sites.

In order to validate the results from GeneClass2, another set of assignment tests were carried out using the software ONCOR (Kalinowski et al. 2008). This software uses conditional maximum likelihood to estimate mixture proportions using the method of Rannala and Mountain (1997). However, both the putative source and sink populations were bootstrapped using 1,000 bootstraps.

It is known that an unbalanced sampling design reduces the power of assignment tests considerably, introducing potential bias in the results (Waples and Gaggiotti 2006). Since the sample sizes between sampling sites of the present study were extremely variable, five replicate random subsets were selected from the sites containing larger sample sizes before performing the assignment tests in GeneClass2 and GenoDive in order to have a balanced sample size across sites. For the adult sampling sites, random subsets from the southeast Tasmanian site and the northern fishing zone of South Australia site were selected from the datasets in order to obtain the same sample sizes as the southern fishing zone site from South Australia (n = 16). For the pueruli sampling sites collected during 2012, a random subset of Bicheno, South Arm and Recherche Bay was selected in order to obtain the same sample size as Cape Jaffa (n = 18). Finally, for the pueruli sampling sites collected during 2013, a random subset of Bicheno and South Arm was selected in order to obtain the same sample size as Cape Jaffa (n = 9). The same data subsets were used in both software. Assignment tests were ran five times on different datasets randomly selected and the results reported correspond to the mean and standard error of all five runs. A one-way analysis of variance (ANOVA) was performed to detect significant differences in the mean percentage of the primary, secondary and tertiary assignments of each site and year. If ANOVA was significant, a Tukey post-hoc

test was used to determine where the differences were. ANOVA and Tukey tests were performed in the R package stats v.3.2.4 (R Core Team 2016).

Additionally, to determine if any combination of adults and pueruli *J. edwardsii* belonged to a putative population, Bayesian clustering of all sampling sites was performed using the program STRUCTURE v.2.3.4 (Pritchard et al. 2000) with the first subset of the data that was used for the assignment tests. An admixture model with no populations selected a priori was used. The model was run 10 times for values of K from 1 to 10 with a burn-in length of 100,000 iterations and 500,000 replicates for each K. STRUCTURE Harvester (Earl and vonHoldt 2012) was used to determine the most likely number of K present in the dataset based on the Evanno's delta K (Pritchard et al. 2000). To corroborate the output of STRUCTURE, a discriminant analysis of principal components (DAPC) (Jombart et al. 2010) was used. The number of possible genetic clusters based on allele frequency data of adult and puerulus sampling sites was determined using the same subsets of data as above. DAPC was performed and results were plotted using the R package adegenet v.2.0.1 (Jombart 2008).

Power analysis

We determined the power of the putatively neutral and outlier SNP markers for detecting genetic structure with the program POWSIM (Ryman and Palm 2006). Effective population size (N_e) was set to 3,000 individuals, according to the findings of Thomas & Bell (2013) and considering 10 generations of drift (t). The generations of drift as well as N_e were used to calculate true F_{ST} values (Ryman and Palm 2006). Since the current version of POWSIM can only analyze 50 loci (Ryman and Palm 2006), only the statistical power of the first 50 neutral SNPs (out of 306) were calculated for the populations collected during 2012 and 2013.

5.4 Results

Sequencing data and preliminary analyses of genetic diversity

Sequencing of five Illumina HiSeq 2500 lanes yielded an average of 2.4 million reads per sample. The reference catalog yielded 5,425 loci. After read alignment and variant calling, a total of 6,236 variable sites were obtained. After selecting one SNP per locus, filtering for MAF, LD and missing data, the final number of SNPs was 314. The PCA plot showed each technical replicate pair was distributed close to each other. This suggests that sequencing samples across multiple lanes did not introduce a large bias in the catalogue building process (Fig. S1, Supporting information).

The SNP outlier identification carried out in LOSITAN produced a panel of eight SNPs under putative positive selection, 85 SNPs under putative balancing selection and 221 neutral SNPs. OutFLANK detected seven SNPs under putative positive selection and 307 neutral SNPs. Six SNPs under putative positive selection were in common between the two outlier identification methods and therefore were excluded from downstream analyses. The final panel of markers was composed of 308 SNPs.

There was low overall genetic differentiation during the two years analysed ($F_{ST,2012}$ = -0.082, n.s. ; $F_{ST,2013}$ = -0.011, n.s.). There was also slight but significant genetic differentiation between adult *J. edwardsii* from each of the sampling sites (Table S2, Supporting information). Pueruli settling into Cape Jaffa during two consecutive years were genetically similar, whereas pueruli settling into South Arm and Bicheno during 2012 differed genetically from those settling within each site during the year 2013 (Table S3, Supporting information).

Adult samples also exhibited slight but significant genetic differentiation against all puerulus samples collected during 2012 (Table 5.2). When analyzing genetic differences between adult *J. edwardsii* and pueruli settling during 2013 there was a different trend to the

found during 2012, suggesting that pueruli source populations vary year-to-year. Pueruli settling into Cape Jaffa and Bicheno were genetically similar to adults settling into southeast Tasmania and distinct to adults sampled in South Australia (Table 5.2). Pueruli settling into South Arm were genetically distinct from all adult sampling sites (Table 5.2).

Table 5.2 Nei's pairwise F_{ST} values between adult and pueruli *J. edwardsii* sampled during years 2012 and 2013 in South Australia and Tasmania. Values in bold are significant at an α=0.05 after a FDR correction

	Cape Jaffa	South Arm	Recherche Bay	Bicheno
Year 2012				
Northern fishing zone SA	**0.0116**	**0.0062**	**0.0092**	**0.0084**
Southern fishing zone SA	**0.0156**	**0.0052**	**0.0226**	**0.0204**
Southeast Tasmania	**0.0053**	**0.0030**	**0.0045**	**0.0034**
Year 2013				
Northern fishing zone SA	**0.0093**	**0.0221**	NA	**0.0087**
Southern fishing zone SA	**0.0097**	**0.0097**	NA	**0.0104**
Southeast Tasmania	0.0016	**0.0082**	NA	0.0024

Assignment to population of origin

For GeneClass2, significant differences between the percentages of puerulus assignment to three adult sampling sites were found at Recherche Bay during the year 2012 (F = 3.935, *P* = 0.0485). Post-hoc comparisons indicated that the primary assignment, the northern fishing zone of South Australia, was significantly different from the tertiary assignment, the southern fishing zone of South Australia (Tukey's *P* = 0.048). No significant differences between the primary, secondary and tertiary assignments were detected for any of the other sampling sites and years (Table 5.3). Results of the assignment tests performed for each random subset are presented in Table S4 (Supporting information).

For ONCOR, significant differences between the percentages of puerulus assignment to three adult sampling sites were found at Bicheno and the South Arm during the year 2012 and

Cape Jaffa during the years 2012 and 2013. Post-hoc comparisons indicated that the primary assignment was significantly different from the second and third assignments in all cases (Table 5.4). Results of the assignments tests performed for ach random subset are presented in Table S5 (Supporting information). It is important to note for both assignment methods the primary assignment was different during both study years.

Additionally, three putative populations were detected by STRUCTURE when assessing the three adults sampling sites and the four pueruli sampling sites collected during the year 2012. Eight putative populations were identified when assessing the three adults sampling sites and the three pueruli sampling sites collected during the year 2013. However, there was no clear differentiation of individuals into separate clusters and therefore there was a lack of population structure. Similarly, there was no evidence of the presence of groups of individuals with more similar genetic identity within the dataset (DAPC) and therefore this tool could not assign pueruli to putative population of origin (Figures S2, S3, Supporting information).

Power analysis

For the populations collected during the year 2012, the overall power estimate of 50 neutral SNPs from the 1,000 simulated runs was 0.2970 for Chi-square and 0.2890 for the Fisher approach, based at a significance level of $\alpha = 0.05$. For the populations collected during the year 2013, the overall power estimate of 50 neutral SNPs from the 1,000 simulated runs was 0.1940 for Chi-square and 0.1970 for the Fisher approach, based at a significance level of $\alpha = 0.05$.

Table 5.3 Most likely population of origin of pueruli sampled during the years 2012 and 2013 in South Australia and Tasmania. Columns three to eight show the most likely populations and the mean (±SE) of their relative scores calculated from five runs of assignment tests performed on GeneClass2. Values in bold indicate significant differences between primary, secondary and/or tertiary assignment

Puerulus sampling site	Settlement year	Primary assignment	Mean score % (±SE)	Secondary assignment	Mean score % (±SE)	Tertiary assignment	Mean score % (±SE)
Cape Jaffa	2012	NFZ of SA	44.521 (±20.386)	SE Tasmania	35.116 (±21.338)	SFZ of SA	20.362 (±17.334)
	2013	SFZ of SA	54.661 (±22.763)	SE Tasmania	45.098 (±22.852)	NFZ of SA	0.241 (±0.149)
Recherche Bay	2012	**NFZ of SA**	**56.814 (±18.278)**	SE Tasmania	43.089 (±18.262)	**SFZ of SA**	**0.097 (±0.060)**
South Arm	2012	SE Tasmania	36.661 (±19.026)	NFZ of SA	35.296 (±21.586)	SFZ of SA	30.043 (±18.268)
	2013	SFZ of SA	42.871 (±23.298)	SE Tasmania	37.346 (±22.879)	NFZ of SA	19.827 (±18.660)
Bicheno	2012	NFZ of SA	60.992 (±19.063)	SE Tasmania	39.007 (±19.063)	SFZ of SA	0.000
	2013	SFZ of SA	55.434 (±19.122)	SE Tasmania	33.217 (±20.589)	NFZ of SA	11.350 (±11.285)

SFZ of SA is the southern fishing zone of South Australia; NFZ of SA is the northern fishing zone of South Australia; SE Tasmania is southeast Tasmania

Table 5.4 Most likely population of origin of pueruli sampled during the years 2012 and 2013 in South Australia and Tasmania. Columns three to eight show the most likely populations and the mean (±SE) of their relative scores calculated from five runs of assignment tests performed on ONCOR

Puerulus sampling site	Settlement year	Primary assignment	Mean score % (±SE)	Secondary assignment	Mean score % (±SE)	Tertiary assignment	Mean score % (±SE)
Cape Jaffa	2012	NFZ of SA	48.194 (±0.027)**	SE Tasmania	26.406 (±0.065)*	SFZ of SA	25.396 (±0.043)*
	2013	SE Tasmania	48.442 (±0.102)**	SFZ of SA	48.108 (±0.083)*	NFZ of SA	3.450 (±0.021)*
Recherche Bay	2012	NFZ of SA	37.828 (±0.045)*	SE Tasmania	33.724 (±0.045)*	SFZ of SA	28.444 (±0.024)*
South Arm	2012	SFZ of SA	44.108 (±0.044)**	NFZ of SA	31.566 (±0.029)*	SE Tasmania	24.328 (±0.016)*
	2013	SFZ of SA	37.376 (±0.040)*	SE Tasmania	33.176 (±0.073)*	NFZ of SA	29.450 (±0.075)*
Bicheno	2012	NFZ of SA	56.942 (±0.095)**	SE Tasmania	34.376 (±0.077)	SFZ of SA	8.682 (±0.040)**
	2013	SFZ of SA	49.372 (±0.111)*	SE Tasmania	27.458 (±0.108)*	NFZ of SA	23.172 (±0.072)*

SFZ of SA is the southern fishing zone of South Australia; NFZ of SA is the northern fishing zone of South Australia; SE Tasmania is southeast Tasmania
* indicates non significant differences between mean assignment scores; ** indicates significant differences between mean assignment scores

5.5 Discussion

The present study constitutes a preliminary investigation of directionality of *J. edwardsii*'s larval migration in southeast Australia. Three sampling sites of adult *J. edwardsii* and four puerulus sampling sites were analyzed. From the adult sampling sites, the South Australian sites were identified as likely populations of origin of pueruli settling into South Australia and Tasmania, in a pattern consistent over two consecutive years. Self-recruitment was also evidenced in Tasmania by the important contribution from the sampling site in southeast Tasmania into all three puerulus collection sites. The results suggested a variation in source populations between years in the Tasmanian sites; indicating ephemeral temporal genetic divergence in *J. edwardsii* recently settled pueruli and corroborating the findings of chapter 3 of this thesis. However, limitations in the sampling design hinder reaching definite conclusions about how these findings support transport pathways predicted for this species in hydrodynamic modeling (Bruce et al. 2007).

Assignment of pueruli to population of origin

The high assignment scores that the adult *J. edwardsii* sampling sites exhibited in the present study are due to the fact that assignment tests assume that sampled populations represent all the populations of a species (Beerli 2004). Nevertheless, the assignment tests identified all three adult sampling sites as likely populations of origin of pueruli settling into Cape Jaffa, South Arm and Bicheno during the years 2012 and 2013. The southern fishing zone of South Australia has been proposed as a contributor of recruits into the Tasmanian fishery due to the high productivity of the area (Linnane et al. 2010c) and the flow of currents displacing larvae south-eastwardly (Bruce et al. 2007). Larval dispersal simulations suggest that larvae originated in

Spencer (close to the northern fishing zone of South Australia) can successfully recruit into all regions across the Great Australian Bight and North-eastern Tasmania (Bruce et al. 2007), which encompasses the site of Bicheno sampled in the present study. Therefore, the results found herein suggest that *J. edwardsii* in southeast Australia experiences a combination of genetic exchange and self-recruitment between and within regions, respectively.

The genetic divergence found between samples collected during 2012 and 2013 in Bicheno and South Arm are in support of chapter 3 findings, which identified chaotic genetic patchiness in recently settled pueruli. Similarly Bruce et al.'s (2007) model assessed during a period of ten years also predicted year-to-year differences in larval trajectories. Chapters 3 and 4 of this thesis found genetic divergence between settling pueruli when assessing regions of the genome under selection. However, the loci under putative positive selection could not be attributed to any gene, therefore we can only hypothesize that the variability in settlement observed in the present study could also due to natural selection driven by different environmental factors each year.

Even though the present study supports the larval dispersal simulations of Bruce et al. (2007), it is important to acknowledge some limitations inherent to the assignment tests that could introduce bias into these results. Correct assignment of individuals is highly dependent on the level of genetic differentiation between populations (Rosenberg et al. 2003; Waples and Gaggiotti 2006). High level of genetic differentiation and low migration rates provide accurate estimates of recent migration (Wilson and Rannala 2003). The overall genetic differentiation found herein was negligible during the two years analyzed ($F_{ST,2012}$ = -0.082; $F_{ST,2013}$ = -0.011) and although most pairwise F_{ST} values differed significantly, they also showed small

levels of divergence. The low genetic differentiation between the populations sampled in the present study could explain the failure to determine the most likely population of origin of puerulus among the three adult sampling sites. Each of the five subsets of data analyzed exhibited large differences in the assignment scores for each putative source population (Table S4, Supporting information) and the ANOVA displayed non-significant differences among them. This indicates that the assignment tests were not consistently identifying putative source populations and the results should be interpreted with caution.

Additionally, assignment tests require sampling of all populations across a species range in order to have an unbiased detection of migration rates (Slatkin 2004). This is virtually impossible to do in species that extend over long geographical areas (Slatkin 2004). It is for this reason that continually distributed populations with no population boundaries are challenging to any assignment method (Waples and Gaggiotti 2006). The bias introduced by not sampling all populations depends on the level of migration rate from the unknown populations (Beerli 2004). This study only included three sampling sites, and although they were located in areas suggested as important sources of larvae, they include a very narrow representation of genotypes of *J. edwardsii*. Also, based on the low and non-significant overall F_{ST} values, we can assume that there is high immigration between all populations and that there will be high bias in the present results. Including sites along the Tasmanian west and south coasts as well as the state of Victoria may have reduced the bias in the assignment tests and allowed determining with more accuracy whether South Australia is a source of recruits to Tasmania.

The small sampling sizes could have also introduced bias into the analysis. However, simulated genetic data to test Rannala and Mountain's (1997) method

suggest that percentage of correct assignment is more dependent on the type of marker used than on sample size (Waples and Gaggiotti 2006). Despite assertions that microsatellites are more informative markers due to a larger number of alleles (Rosenberg et al. 2003), SNP markers have been suggested to be as efficient at assigning individuals to their population of origin (Smith and Seeb 2008).

Finally, the adult samples comprised individuals of multiple ages. This was evidenced by the significant genetic divergence between years within Tasmanian puerulus sampling sites, suggesting that pueruli source populations vary year-to-year. Knowledge of adult specimens age would be needed in order to determine their year of settlement. Even when protocols for age estimation of decapods have been recently developed (Kilada et al. 2012; Leland et al. 2015), there is no conclusive evidence on age estimation of *J. edwardsii*. In the absence of knowledge of adult specimens age, any two adults of similar carapace length could have settled one or two years apart. This implies that adult samples from any one site were potentially sourced from different populations. This limitation is impossible to overcome in the study species and it would be present even if sampling sites and sizes were increased.

Despite the limitations, the results of the present study could have important implications for the *J. edwardsii* fishery. The results suggest that *J. edwardsii* larvae cross state boundaries and that South Australia is an important larval source for southeast Tasmania. This suggests that the dispersal model developed by Bruce et al. (2007) should be taken into account when making managerial decisions. The egg production in the southern fishing zone of South Australia during the 2013 fishery status assessment was at 9% of its unfished level (Linnane et al. 2014b). This is well below the recommended egg production set by the commonwealth fisheries harvest strategy, which is 20% of the unfished virgin level (Australia Department of

Agriculture, Fisheries and Forestry 2007). Despite egg production being very low, the total allowable commercial catch (TACC) was retained at the same level for the 2014-2015 season (Linnane et al. 2014b). If this region provides a high percentage of recruits into eastern areas, then there is imminent need of rebuilding this stock.

5.6 Conclusions

Based on the present findings it is not possible to determine the most likely population of origin of puerulus settling into southeast Tasmania and South Australia. The low levels of genetic differentiation and the fact that the South Australian and Tasmanian adult sampling sites contributed equally to the analyzed pueruli suggest the genetic admixture of *J. edwardsii* in southeast Australia. This highlights the importance of joint management between Australian states in order to rebuild depleted stocks that act as source of recruits. However, due to sampling design limitations, results should be interpreted with caution and taken as a preliminary assessment of the directionality of migration of *J. edwardsii* in Australia.

5.7 Acknowledgements

Funding for this research was provided by an Australian Research Council Linkage Project grant (Project No. LP120200164) and Australian Research Council Discovery Project grant (Project No. DP150101491). We would like to thank Peter Hawthorne (South Australian Research and Development Institute, SARDI) for puerulus and adult collections from South Australia. Thanks to Gary Carlos (University of Tasmania), Colin Fry (University of Tasmania) and Andrew Kent for field assistance and adult sample collection in Australia. We thank Ivan Hinojosa, Ruari Colquhoun, David Fallon, Kylie Cahill, Graeme Ewing and Sarah Pyke for puerulus collections in Tasmania. Thanks to Brooke Whitelaw and Daniel Laird for laboratory assistance. Special thanks to Karen J. Miller who contributed to the original project idea.

5.8 Supplementary information

Table S1 Nei's pairwise F_{ST} values between three sampling sites of adult lobsters in Tasmania. F_{ST} values were non significant at an α=0.05

	Bruny Island	Taroona
Betsey Island	-0.0018	0.0001
Bruny Island		0.0025

Table S2 Nei's pairwise F_{ST} values between adult sampling sites in South Australia and Tasmania. Values in bold are significant at an α=0.05 after a FDR correction

	Southern fishing zone SA	Southeast Tasmania
Northern fishing zone SA	**0.0136**	**0.0056**
Southern fishing zone SA	-	**0.0066**

Table S3 Nei's pairwise F_{ST} values between pueruli sampling sites in South Australia and Tasmania during the years 2012 and 2013. Values in bold are significant at an α=0.05 after a FDR correction

	Cape Jaffa 2013	South Arm 2013	Bicheno 2013
Cape Jaffa 2012	0.0012	-	-
South Arm 2012	-	**0.0047**	-
Bicheno 2012	-	-	**0.0038**

Table S4 Most likely population of origin of pueruli sampled during the years 2012 and 2013 in South Australia and Tasmania using five random subsets of data in GeneClass2. Columns three to eight show the most likely populations and their relative scores

Puerulus sampling site	Settlement year	Primary assignment	Score %	Secondary assignment	Score %	Tertiary assignment	Score %
Subset 1							
Cape Jaffa	2012	SFZ of SA	89.567	NFZ of SA	9.850	SE Tasmania	0.583
	2013	SFZ of SA	99.996	NFZ of SA	0.002	SE Tasmania	0.001
Recherche Bay	2012	SE Tasmania	69.623	NFZ of SA	30.088	SFZ of SA	0.288
South Arm	2012	NFZ of SA	89.493	SFZ of SA	10.319	SE Tasmania	0.189
	2013	SFZ of SA	99.891	NFZ of SA	0.109	SE Tasmania	0.000
Bicheno	2012	NFZ of SA	99.998	SE Tasmania	0.002	SFZ of SA	0.000
	2013	SFZ of SA	100	NFZ of SA	0.000	SE Tasmania	0.000
Subset 2							
Cape Jaffa	2012	NFZ of SA	94.208	SFZ of SA	5.442	SE Tasmania	0.370
	2013	SFZ of SA	74.016	SE Tasmania	25.471	NFZ of SA	0.513
Recherche Bay	2012	NFZ of SA	85.860	SE Tasmania	14.140	SFZ of SA	0.000
South Arm	2012	SE Tasmania	68.067	SFZ of SA	31.921	NFZ of SA	0.012
	2013	SE Tasmania	91.053	SFZ of SA	4.545	NFZ of SA	4.402
Bicheno	2012	SE Tasmania	76.664	NFZ of SA	23.335	SFZ of SA	0.000
	2013	NFZ of SA	56.491	SFZ of SA	43.392	SE Tasmania	0.118

SFZ of SA is the southern fishing zone of South Australia; NFZ of SA is the northern fishing zone of South Australia; SE Tasmania is southeast Tasmania

Table S4 Continued

Puerulus sampling site	Settlement year	Primary assignment	Score %	Secondary assignment	Score %	Tertiary assignment	Score %
Subset 3							
Cape Jaffa	2012	SE Tasmania	90.779	NFZ of SA	8.719	SFZ of SA	0.502
	2013	SE Tasmania	99.965	SFZ of SA	0.035	NFZ of SA	0.000
Recherche Bay	2012	NFZ of SA	67.405	SE Tasmania	32.404	SFZ of SA	0.192
South Arm	2012	SE Tasmania	91.873	SFZ of SA	7.984	NFZ of SA	0.143
	2013	NFZ of SA	94.391	SFZ of SA	5.604	SE Tasmania	0.004
Bicheno	2012	SE Tasmania	90.750	NFZ of SA	9.250	SFZ of SA	0.000
	2013	SFZ of SA	99.979	NFZ of SA	0.020	SE Tasmania	0.000
Subset 4							
Cape Jaffa	2012	SE Tasmania	83.848	NFZ of SA	15.273	SFZ of SA	0.879
	2013	SFZ of SA	99.860	SFZ of SA	0.139	NFZ of SA	0.001
Recherche Bay	2012	SE Tasmania	99.277	NFZ of SA	0.719	SFZ of SA	0.004
South Arm	2012	NFZ of SA	86.830	SE Tasmania	13.170	SFZ of SA	0.000
	2013	SFZ of SA	95.668	SFZ of SA	4.332	NFZ of SA	0.000
Bicheno	2012	NFZ of SA	72.379	SE Tasmania	27.621	SFZ of SA	0.000
	2013	SE Tasmania	72.428	SFZ of SA	27.566	NFZ of SA	0.006

SFZ of SA is the southern fishing zone of South Australia; NFZ of SA is the northern fishing zone of South Australia; SE Tasmania is southeast Tasmania

Table S4 Continued

Puerulus sampling site	Settlement year	Primary assignment	Score %	Secondary assignment	Score %	Tertiary assignment	Score %
Subset 5							
Cape Jaffa	2012	NFZ of SA	94.557	SFZ of SA	5.442	SE Tasmania	0.000
	2013	SFZ of SA	99.119	NFZ of SA	0.687	SE Tasmania	0.194
Recherche Bay	2012	NFZ of SA	99.998	SFZ of SA	0.002	SE Tasmania	0.000
South Arm	2012	SFZ of SA	99.992	SE Tasmania	0.008	NFZ of SA	0.000
	2013	SFZ of SA	99.985	NFZ of SA	0.009	SE Tasmania	0.006
Bicheno	2012	NFZ of SA	100	SFZ of SA	0.000	SE Tasmania	0.000
	2013	SE Tasmania	93.537	SFZ of SA	6.232	NFZ of SA	0.231

SFZ of SA is the southern fishing zone of South Australia; NFZ of SA is the northern fishing zone of South Australia; SE Tasmania is southeast Tasmania

Table S5 Most likely population of origin of pueruli sampled during the years 2012 and 2013 in South Australia and Tasmania using five random subsets of data in ONCOR. Columns three to eight show the most likely populations and their relative scores

Puerulus sampling site	Settlement year	Primary assignment	Score %	Secondary assignment	Score %	Tertiary assignment	Score %
Subset 1							
Cape Jaffa	2012	NFZ of SA	40.15	SE Tasmania	37.21	SFZ of SA	22.63
	2013	SFZ of SA	72.32	SE Tasmania	19.68	NFZ of SA	0.08
Recherche Bay	2012	NFZ of SA	45.82	SFZ of SA	30.87	SE Tasmania	23.30
South Arm	2012	SFZ of SA	47.48	NFZ of SA	29.20	SE Tasmania	23.32
	2013	SFZ of SA	44.57	NFZ of SA	41.56	SE Tasmania	13.87
Bicheno	2012	NFZ of SA	52.36	SE Tasmania	27.62	SFZ of SA	20.02
	2013	SFZ of SA	85.88	NFZ of SA	14.12	SE Tasmania	0.00
Subset 2							
Cape Jaffa	2012	NFZ of SA	51.22	SFZ of SA	25.00	SE Tasmania	23.78
	2013	SE Tasmania	55.23	SFZ of SA	44.74	NFZ of SA	0.00
Recherche Bay	2012	SE Tasmania	44.54	NFZ of SA	33.87	SFZ of SA	21.59
South Arm	2012	SFZ of SA	46.64	NFZ of SA	29.70	SE Tasmania	23.67
	2013	SE Tasmania	39.36	NFZ of SA	37.52	SFZ of SA	23.13
Bicheno	2012	SE Tasmania	47.74	NFZ of SA	35.71	SFZ of SA	16.55
	2013	NFZ of SA	50.14	SFZ of SA	28.31	SE Tasmania	21.56

SFZ of SA is the southern fishing zone of South Australia; NFZ of SA is the northern fishing zone of South Australia; SE Tasmania is southeast Tasmania

Table S5 Continued

Puerulus sampling site	Settlement year	Primary assignment	Score %	Secondary assignment	Score %	Tertiary assignment	Score %
Subset 3							
Cape Jaffa	2012	NFZ of SA	47.64	SE Tasmania	36.71	SFZ of SA	15.64
	2013	SE Tasmania	63.53	SFZ of SA	36.45	NFZ of SA	0.00
Recherche Bay	2012	SFZ of SA	35.67	SE Tasmania	33.30	NFZ of SA	31.03
South Arm	2012	NFZ of SA	36.28	SFZ of SA	35.43	SE Tasmania	28.29
	2013	NFZ of SA	40.71	SFZ of SA	40.04	SE Tasmania	19.25
Bicheno	2012	NFZ of SA	49.08	SE Tasmania	48.85	SFZ of SA	2.01
	2013	SFZ of SA	63.61	NFZ of SA	25.57	SE Tasmania	10.82
Subset 4							
Cape Jaffa	2012	NFZ of SA	45.67	SE Tasmania	30.97	SFZ of SA	23.36
	2013	SE Tasmania	73.77	SFZ of SA	26.22	NFZ of SA	0.00
Recherche Bay	2012	SE Tasmania	43.34	SFZ of SA	29.22	NFZ of SA	27.44
South Arm	2012	NFZ of SA	39.54	SFZ of SA	33.52	SE Tasmania	26.94
	2013	SE Tasmania	53.71	SFZ of SA	44.61	NFZ of SA	16.80
Bicheno	2012	NFZ of SA	55.07	SE Tasmania	40.17	SFZ of SA	4.76
	2013	SE Tasmania	51.72	SFZ of SA	38.70	NFZ of SA	9.58

SFZ of SA is the southern fishing zone of South Australia; NFZ of SA is the northern fishing zone of South Australia; SE Tasmania is southeast Tasmania

Table S5 Continued

Puerulus sampling site	Settlement year	Primary assignment	Score %	Secondary assignment	Score %	Tertiary assignment	Score %
Subset 5							
Cape Jaffa	2012	NFZ of SA	56.29	SFZ of SA	41.57	SE Tasmania	2.14
	2013	SFZ of SA	60.81	SE Tasmania	30.00	NFZ of SA	9.19
Recherche Bay	2012	NFZ of SA	50.98	SFZ of SA	24.87	SE Tasmania	24.14
South Arm	2012	SFZ of SA	57.47	NFZ of SA	23.11	SE Tasmania	19.42
	2013	SE Tasmania	39.69	SFZ of SA	34.53	NFZ of SA	25.78
Bicheno	2012	NFZ of SA	92.49	SE Tasmania	0.75	SFZ of SA	0.00
	2013	SE Tasmania	53.19	SFZ of SA	30.36	NFZ of SA	16.45

SFZ of SA is the southern fishing zone of South Australia; NFZ of SA is the northern fishing zone of South Australia; SE Tasmania is southeast Tasmania

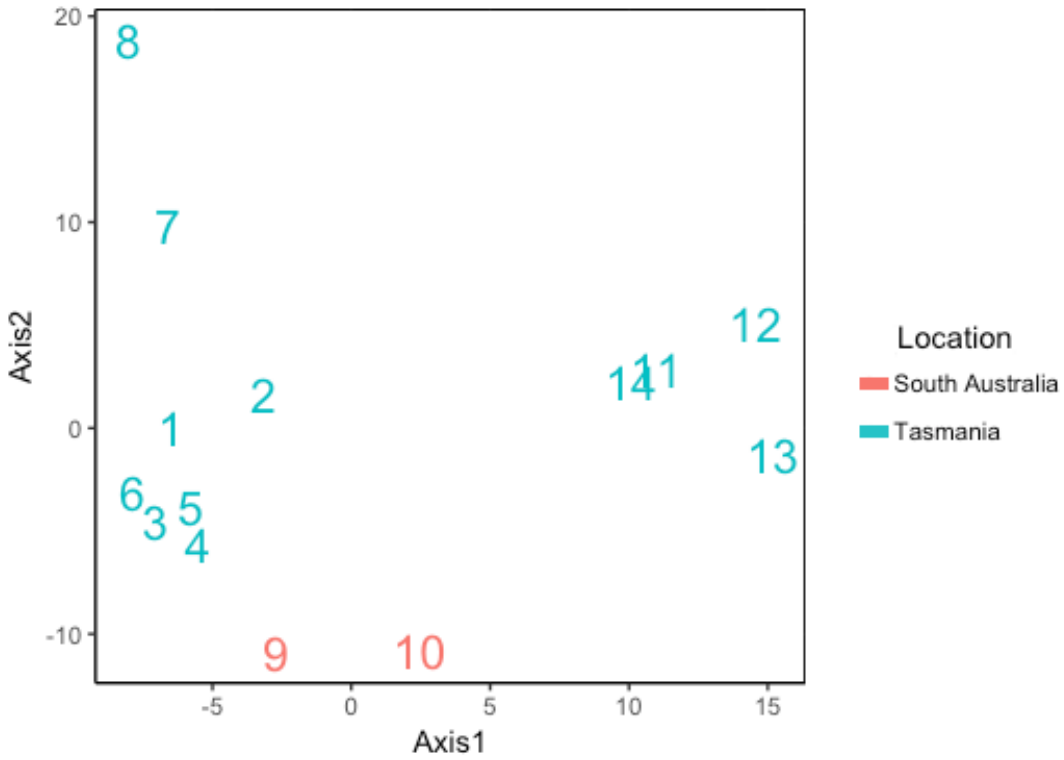


Fig. S1 Principal component analysis based on allele frequencies of technical replicates from Bicheno (blue numbers) and Cape Jaffa (red numbers). Consecutive numbers from 1 to 10 belong to each intralibrary technical replicate pair (e.g. 1&2, 3&4, etc. represent replicate pair). Samples 11 to 14 constitute interlibrary technical replicates

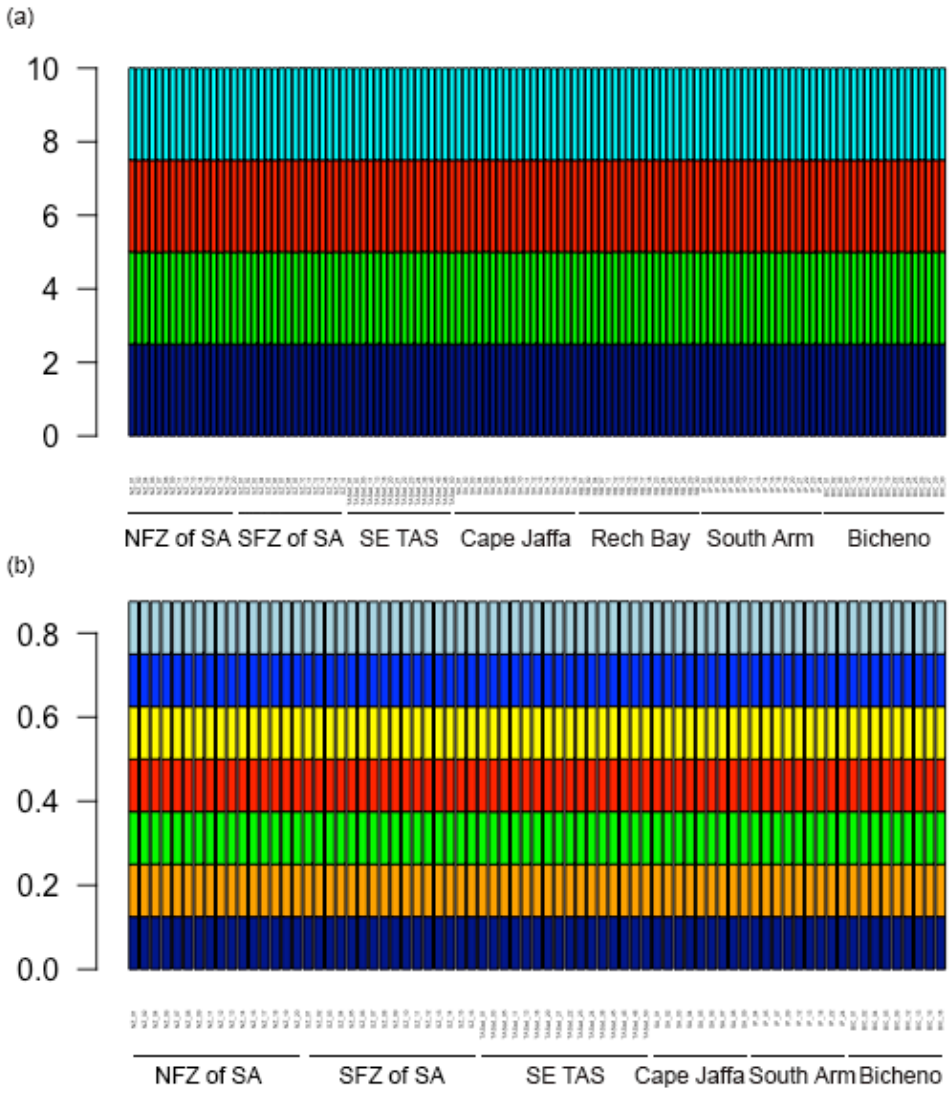


Fig. S2 Genetic differentiation between adult and puerulus sampling sites in South Australia and Tasmania during 2012 (a) and 2013 (b) using Bayesian clustering. NFZ of SA is the northern fishing zone of South Australia; SFZ of SA is the southern fishing zone of South Australia; SE TAS is southeast Tasmania; Rech Bay is Recherche Bay

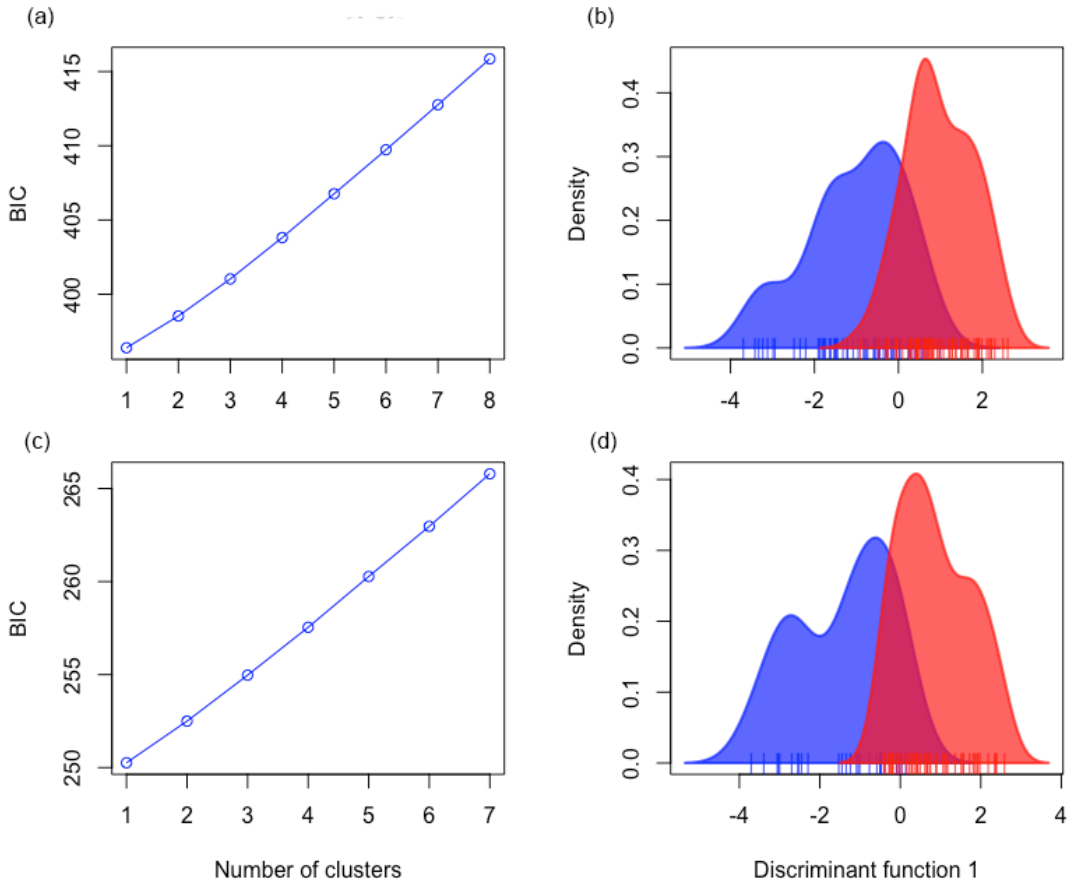


Fig. S3 Genetic differentiation between adult and puerulus sampling sites in South Australia and Tasmania during 2012 (a, b) and 2013 (c, d) using a discriminant analysis of principal components. The left panels show the Bayesian index criteria (BIC) plots used to define the most likely number of groups present. The right panels show the first principal component resulting from the discriminant analysis of principal components

Chapter 6: General Discussion

6.1 Thesis overview

This thesis demonstrated genetic divergence between *Jasus edwardsii* populations in Australia and New Zealand, separated by 1,000's of kilometres, but genetic admixture within Australia. In addition, high variability in recruitment abundance of *J. edwardsii* was also reflected in the genetic identity of recruits. Population structure in *J. edwardsii* recruits did not follow the isolation by distance model, but was rather chaotic, as is common in highly fecund species with protracted pelagic larval durations (PLDs). Chaotic genetic patchiness was observed at both medium (100's km) and small (10's km) spatial scales as well as between years and within a recruitment season.

Adaptive processes shaped population structure of *J. edwardsii* at different geographical scales. Adult *J. edwardsii* exhibited local adaptation at a broad geographical scale (chapter 2). This implied reduced larval exchange between Australia and New Zealand, and is in line with the hypothesis of the Tasman Sea acting as a barrier to dispersal (Thomas and Bell 2013). Local adaptation within each country was evidenced by differences in loci under putative selection. Investigation of *J. edwardsii* recruits in two sites in southeast Australia further demonstrated weak natural selection acting consistently across time and space (chapter 3). This suggested that during the first weeks of benthic life, *J. edwardsii* pueruli are subject to less natural selection than during the PLD, providing support for the environment driving the observed population structure in this species.

On a finer scale, there were phenotypic differences in *J. edwardsii* recruits arriving at different sites along southeast Tasmania and during three consecutive months during the winter recruitment peak. This suggested dispersal history was a cause of

recruitment variability (chapter 4). It is possible that larval *J. edwardsii* hatching during the same event could be transported offshore (Booth and Phillips 1994) and therefore experience their PLD as an assemblage. Genetic differences at neutral loci in one of the study sites demonstrated chaotic genetic patchiness, adding strength to the hypothesis of collective dispersal despite the long PLD of this species.

Finally, a preliminary assessment of the source of recruits settling into South Australia and Tasmania also evidenced genetic admixture of *J. edwardsii* in southeast Australia (chapter 5). These findings were in line with predictions of a larval advection simulation study in southern Australia (Bruce et al. 2007). However, the low level of genetic divergence between populations and the insufficient number of sampling sites hindered detection of the most likely population of origin of recruits. All the adult populations included in the analysis were equally likely of contributing to the recruits in one site in South Australia and three sites in Tasmania.

6.2 Implications of the findings for fisheries management of this species

This thesis supported the hypothesis of a panmictic population of *J. edwardsii* in southeast Australia. This means that populations are linked genetically throughout their range. *J. edwardsii* is fished both commercially and recreationally in four states, Western Australia, South Australia, Victoria and Tasmania. Although there is evidence of high larval exchange between jurisdictions (Ovenden et al. 1992; Bruce et al. 2007) egg production in each state is managed independently. Consequently, if low levels of egg production in one jurisdiction had a negative impact on recruitment elsewhere then it could be beyond the power of the state government to respond. The need for broad scale management of egg production thus relies on cooperative management across jurisdictions. A specific concern for the fishery that has been apparent for many years is that the depleted egg production in the South Australian

southern zone increases the risk of periods of low recruitment occurring across the wider coast including in others states (i.e. Victoria and Tasmania; Bruce et al. 2007).

The reference points taken into consideration to set the total allowable commercial catch (TACC) in South Australia's fishery are catch per unit effort (CPUE) in the previous fishing season and pre-recruit index (PRI) (Linnane et al. 2014b). However, there is no performance indicator associated with preventing recruitment overfishing. Such performance indicator needs to be a direct measure or proxy for reproductive output, such as egg production. The stock assessment model of the southern fishing zone of South Australia for the 2013 season estimated that egg production was at 9% of virgin levels (Linnane et al. 2014b). This was much lower than the recommended biomass limit reference point set by the commonwealth harvest strategy policy of 20% (Australia Department of Agriculture, Fisheries and Forestry 2007). Despite this, the TACC was maintained at the same level during the following fishing season (Linnane et al. 2014b). The temporal variability in the genetic identity of recruits found in chapters 3 and 4 suggests that adult lobsters from different zones could vary in their contribution to the population over time. If the egg production in some areas is below healthy levels, keeping the same TACC could exert high fishing pressure on populations with low levels of egg production. To avoid this, management efforts such as lowering the TACC should be put in place when egg production falls below a minimum limit.

Managing egg production at small spatial scales is also important in order to ensure that the spawning stocks across all jurisdictions will be maintained at healthy levels. Catch quotas in Tasmania are currently set based on a bioeconomic projection model that evaluates different management strategies against biomass, egg production, CPUE-based limits and target reference points (Plagányi et al. 2017). The limit reference point of virgin egg production is set to an overall statewide level of

20% (Australia Department of Agriculture, Fisheries and Forestry 2007). Yet egg production across fishing zones varies considerably, from 13% in the north west of the state (fishing zone 5) up to 81% in southeast Tasmania (fishing zone 8) (Gardner et al. 2015a). The chaotic genetic patchiness found in this thesis suggests the need of setting a reference point for egg production for each individual fishing zone. Restoring egg production through better management of the spatial patterns in the fishery would reduce the risk of years with low puerulus settlement and low fishery productivity.

6.3 Limitations

The main limitation of the present study was the relatively low number of molecular markers generated and the modest sample sizes used in all data chapters. These were due to the large amount of missing data per sample after SNP calling and resulted in a very high percentage of loci as well as samples being discarded for downstream analyses. ddRADseq is known for producing large amounts of missing data due to wet-lab procedures and bioinformatics (Andrews and Luikart 2014). Sequencing a smaller amount of samples within each ddRADseq library (or HiSeq lane) could produce a higher number of sample/reads available, however this would increase the sequencing costs.

Reducing the fragment size of the sequenced libraries could have also increased the marker depth and avoided discarding loci due to low representation (Stolle and Moritz 2013). Given that some samples sequenced for the present thesis were used in more than one data chapter, there was the need of maximizing the overlapping regions between sequenced ddRADseq libraries. Therefore, reducing the fragment size could have resulted in non-overlapping regions between ddRADseq libraries. This would

have resulted in loci that were not shared between libraries, which would have also been discarded, reducing the overall lower number of SNPs.

The large genome of *J. edwardsii* and the lack of a reference genome possibly contributed to the amount of missing data. This is because large genomes possess repetitive regions of DNA which result in large amount of paralogous loci (Li et al. 2003). The pipeline used to generate the reference catalogue of loci (rad-loci pipeline) maximizes the removal of paralogous sequences (explained in detail in chapter 3) and by doing this, a large amount of reads are discarded from the dataset. Alignment of the sequenced reads to a reference genome would have enabled identifying paralogous regions (Andrews and Luikart 2014), allowing a more efficient use of the sequenced reads.

Despite the relatively low number of molecular markers, the power analysis implemented in chapter 4 determined that a small subset of SNPs (50 SNPs out of a total of 603 SNPs) had approximately 40% power to detect differences between populations. This subject was particularly important to address given that chaotic genetic patchiness can arise due to a sampling artifact of randomly sub-sampling a small proportion of individuals out of a very large population. For chapter 5, the power analyses detected a lesser power of the SNPs to detect genetic differences among populations. For the populations sampled during the year 2012, 50 SNPs only had 30% power to detect population differentiation; whereas for the populations sampling during 2013, the power of 50 SNPs was reduced to 20%. However, 50 SNPs constitute 1/6 of the total number of markers included in the genetic analyses (n = 306). Therefore, it is expected that increasing the number of markers would give more than 50% power to the genetic analyses. The limitations of the assignment tests were therefore related to the sampling design, rather than to the low number of markers.

The bias introduced by the non-sampled, "ghost", populations obscured the results of the assignment tests (Waples and Gaggiotti 2006) in chapter 5. Increasing the number and spatial distribution of sampling sites could have allowed more definite conclusions to be reached about the directionality of the migration. Including more sampling sites as putative source populations in the assignment tests would have also allowed testing the alternative hypothesis of the existence of other areas outside South Australia as the main contributors of pueruli into Tasmania.

6.4 Future research directions

Chapter 5 of this thesis provided a preliminary assignment of puerulus into putative population of origin in order to test the findings of the larval transport simulations by Bruce et al. (2007). However, there was an insufficient number and distribution of sampling sites for adult *J. edwardsii,* which prevented reaching conclusive results. Therefore, the source of recruits settling into Tasmania remains to be tested. Extending the number and distribution of sampling sites would help obtaining accurate results about the level of inter-State dispersal and within-State self-recruitment. This could be achieved by including adult lobster samples from populations along the west coast of Tasmania, Victoria and South Australia, as well as adult populations from the same sites where the pueruli samples were collected. Furthermore, coupling genetic data to oceanographic modeling would help to measure the effect of ocean advection on larval transport and dispersal of *J. edwardsii*.

Since year-to-year variability in pueruli sources has been already hypothesized (Bruce et al. 2007) and confirmed in chapter 5, studying adult lobsters of different cohorts or size classes is needed in order to reach more accurate conclusions about source and sink populations. Determining the age of analyzed *J. edwardsii* would provide information on adult age classes. Age estimation protocols have been

proposed for other decapod species using gastric ossicles (Kilada et al. 2012; Leland et al. 2015). As part of this thesis I conducted a pilot study ageing adult and puerulus *J. edwardsii*. However, the results were inconclusive due to difficulties with the ageing technique and therefore were not included in this thesis. Determining and validating age increments in *J. edwardsii* would open new avenues of research. This would not only help to address the directionality of migration with more accuracy, but aid fisheries stock assessment by producing age frequency data.

To determine whether the temporal chaotic genetic patchiness (chapter 3) was caused by sweepstakes reproductive success (SRS) it would be helpful to analyze pueruli collected over a longer period of time. In chapter 3 there was no evidence of SRS based on the following: (1) no significant differences between mean genetic diversity of pueruli across years, (2) the lack of genetic cohorts present in pueruli yearly collections and (3) an inability to estimate effective population size (Ne) and therefore to determine whether it varied through time. However, the high fecundity of *J. edwardsii* (Green et al. 2009), its protracted PLD of up to 24 months (Booth 1994) and its low levels of neutral population structure across southeast Australia (reported in this thesis), makes it a model species to investigate SRS. Another way of evidencing SRS could be by investigating if there is a reduction in the genetic diversity of the recently settled pueruli against that of post-pueruli over multiple sites. Determining whether SRS occurs in this species would have important implications for fisheries management, given that low Ne can reduce the genetic diversity of a population, lead to inbreeding depression and affect the population resilience of the resource (Hedgecock 1994).

Finally, developing more genetic resources such as a reference genome and an annotated transcriptome for the species is of prime importance in order to study natural selection in this species as well as its potential for local adaptation. Chapter 2

of this thesis provided evidence of large-scale local adaptation of *J. edwardsii*, while chapters 3 and 4 provided evidence for weak pre and post-settlement selection. Obtaining larger SNP panels of loci under putative selection and screening them against known genes could help determining potential causes of natural selection and genetic differentiation at different spatial scales.

6.5 Conclusions

The findings of this thesis demonstrated limited larval exchange at a broad spatial scale (1,000's km) between Australia and New Zealand but genetic admixture at a medium geographic scale (100's km) within southeast Australia. The population subdivision at broad geographic scale provided evidence for local adaptation between both countries. Despite the overall absence of population structure in southeast Australia, chaotic genetic patchiness existed in newly settled recruits at a medium and fine geographic scale (10's km) as well as between and within years. Based on phenotypic differences of recruits across sites and months of settlement it was possible to hypothesize collective dispersal as a transport mechanism of *J. edwardsii* during the PLD. The findings of this thesis provide new evidence on the influence of dispersal history and natural selection in determining genetic divergence in *J. edwardsii* recruits and adults. This has important implications for the management of egg production in the *J. edwardsii* fishery.

REFERENCES

ABARE-BRS (2009) Australian fisheries statistics 2009. Canberra

Alho JS, Välimäki K, Merilä J (2010) Rhh: an R extension for estimating multilocus heterozygosity and heterozygosity-heterozygosity correlation: computer program note. Molecular Ecology Resources 10:720–722. doi: 10.1111/j.1755-0998.2010.02830.x

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. Nature Reviews Genetics 11:697–709. doi: 10.1038/nrg2844

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 25:3389–3402

Andrews KR, Luikart G (2014) Recent novel approaches for population genomics data analysis. Molecular Ecology 23:1661–1667

Annala JH, Bycroft BL (1985) Growth rate of juvenile rock lobsters *(Jasus edwardsii)* at Stewart Island, New Zealand. New Zealand Journal of Marine and Freshwater Research 19:445–455. doi: 10.1080/00288330.1985.9516109

Annala JH, McKoy JL, Booth JD, Pike RB (1980) Size at the onset of sexual maturity in female *Jasus edwardsii* (Decapoda: Palinuridae) in New Zealand. New Zealand Journal of Marine and Freshwater Research 14:217–227. doi: 10.1080/00288330.1980.9515864

Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008) LOSITAN: a workbench to detect molecular adaptation based on a Fst-outlier method. BMC Bioinformatics 9:323. doi: 10.1186/1471-2105-9-323

Araneda C, Larraín MA, Hecht B, Narum S (2016) Adaptive genetic variation distinguishes Chilean blue mussels (*Mytilus chilensis*) from different marine environments. Ecology and Evolution 6:3632–3644. doi: 10.1002/ece3.2110

Australia Department of Agriculture, Fisheries and Forestry (2007) Commonwealth fisheries harvest strategy: policy and guidelines. Canberra

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. Proceedings of the Royal Society of London B: Biological Sciences 263:1619–1626

Beerli P (2004) Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. Molecular Ecology 13:827–836. doi: 10.1111/j.1365-294X.2004.02101.x

Beldade P, Mateus ARA, Keller RA (2011) Evolution and molecular mechanisms of adaptive developmental plasticity. Molecular Ecology 20:1347–1363. doi: 10.1111/j.1365-294X.2011.05016.x

Benestan L, Gosselin T, Perrier C, Sainte-Marie B, Rochette R, Bernatchez L (2015) RAD genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species, the American lobster (*Homarus americanus*). Molecular Ecology 24:3299–3315. doi: 10.1111/mec.13245

Bernatchez L, Wellenreuther M, Araneda C, Ashton DT, Barth JMI, Beacham TD, Maes GE, Martinsohn JT, Miller KM, Naish KA, Ovenden JR, Primmer CR, Young Suk H, Therkildsen NO, Withler RE (2017) Harnessing the power of genomics to secure the future of seafood. Trends in Ecology & Evolution 32:665–680. doi: 10.1016/j.tree.2017.06.010

Bohonak AJ (1999) Dispersal, gene flow, and population structure. The Quarterly Review of Biology 74:21–45

Booth JD (1979) Settlement of the rock lobster, *Jasus edwardsii* (Decapoda: Palinuridae), at Castlepoint, New Zealand. New Zealand Journal of Marine and Freshwater Research 13:395–406. doi: 10.1080/00288330.1979.9515816

Booth JD (1994) *Jasus edwardsii* larval recruitment off the east coast of New Zealand. Crustaceana 66:295–317

Booth JD, Kittaka J (1994) Growout of juvenile spiny lobster. In: Phillips BF, Cobb JS, Kittaka J (eds) Spiny lobster management, Blackwell Scientific Publications, Oxford, pp 239–429

Booth JD, Phillips BF (1994) Early life history of spiny lobster. Crustaceana 66:271–294

Booth JD, Tarring SC (1986) Settlement of the red rock lobster, *Jasus edwardsii*, near Gisborne, New Zealand. New Zealand Journal of Marine and Freshwater Research 20:291–297. doi: 10.1080/00288330.1986.9516150

Bruce BD, Griffin DA, Bradford RW (2007) Larval transport and recruitment processes of southern rock lobster. CSIRO Marine Research, Hobart

Butler IV MJ, Herrnkind WF (1997) A test of recruitment limitation and the potential for artificial enhancement of spiny lobster (*Panulirus argus*) populations in Florida. Canadian Journal of Fisheries and Aquatic Sciences 54:452–463

Butler IV MJ, Herrnkind WF, Hunt JH (1997) Factors affecting the recruitment of juvenile Caribbean spiny lobsters dwelling in macroalgae. Bulletin of Marine Science 61:3–19

Butler IV MJ, Paris C, Goldstein J, Matsuda H, Cowen R (2011) Behavior constrains the dispersal of long-lived spiny lobster larvae. Marine Ecology Progress Series 422:223–237. doi: 10.3354/meps08878

Butler MJ, Macdiarmid A, Gnanalingam G (2015) The effect of parental size on spermatophore production, egg quality, fertilization success, and larval characteristics in the Caribbean Spiny lobster, *Panulirus argus*. ICES Journal of Marine Science: Journal du Conseil 72:i115–i123. doi: 10.1093/icesjms/fsv015

Caley MJ, Carr MH, Hixon MA, Hughes TP, Jones GP, Menge BA (1996) Recruitment and the local dynamics of open marine populations. Annual Review of Ecology and Systematics 27:477–500

Candy JR, Campbell NR, Grinnell MH, Beacham TD, Larson WA, Narum SR (2015) Population differentiation determined from putative neutral and divergent adaptive genetic markers in Eulachon (*Thaleichthys pacificus*, Osmeridae), an anadromous Pacific smelt. Molecular Ecology Resources 15:1421–1434. doi: 10.1111/1755-0998.12400

Cano JM, Shikano T, Kuparinen A, Merilä J (2008) Genetic differentiation, effective population size and gene flow in marine fishes: implications for stock management. Journal of Integrative Field Biology 5:1–10

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) Stacks: building and genotyping loci de novo from short-read sequences. G3: Genes|Genomes|Genetics 1:171–182. doi: 10.1534/g3.111.000240

Chandrapavan A, Gardner C, Green BS (2010) Growth rate of adult rock lobsters *Jasus edwardsii* increased through translocation. Fisheries Research 105:244–247. doi: 10.1016/j.fishres.2010.03.014

Chandrapavan A, Gardner C, Green BS, Linnane A, Hobday D (2011) Improving marketability through translocation: a lobster case study from southern Australia. ICES Journal of Marine Science: Journal du Conseil 68:1842–1851. doi: 10.1093/icesjms/fsr128

Chiswell S, Booth J (2008) Sources and sinks of larval settlement in *Jasus edwardsii* around New Zealand: where do larvae come from and where do they go? Marine Ecology Progress Series 354:201–217. doi: 10.3354/meps07217

Chiswell SM, Booth JD (1999) Rock lobster Jasus edwardsii larval retention by the Wairarapa Eddy off New Zealand. Marine Ecology Progress Series 183:227–240

Cobb JS, Booth JD, Clancy M (1997) Recruitment strategies in lobsters and crabs: a comparison. Marine and Freshwater Research 48:797. doi: 10.1071/MF97219

Coltman DW, Pilkington JG, Smith JA, Pemberton JM (1999) Parasite-mediated selection against inbred soay sheep in a free-living, island population. Evolution 53:1259. doi: 10.2307/2640828

Corander J, Majander KK, Cheng L, Merilä J (2013) High degree of cryptic population differentiation in the Baltic Sea herring *Clupea harengus*. Molecular Ecology 22:2931–2940. doi: 10.1111/mec.12174

Cowen RK, Lwiza KMM, Sponaugle S, Paris CB, Olson DB (2000) Connectivity of marine populations: open or closed? Science 287:857–859. doi: 10.1126/science.287.5454.857

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST (2011) The variant call format and VCFtools. Bioinformatics 27:2156–2158

Deiana AM, Cau A, Coluccia E, Cannas R, Milia A, Salvadori S, Libertini A (1999) Genome size and AT-DNA content in thirteen species of decapoda. In: Schram R, von Vaupel Klein JC (eds) Crustaceans and the biodiversity crisis, Koninklijke Brill NV, Leiden, The Netherlands, pp 981–985

DeWitt TJ, Sih A, Wilson DS (1998) Costs and limits of phenotypic plasticity. Trends in Ecology & Evolution 13:77–81

Dingeldein AL, White JW (2016) Larval traits carry over to affect post-settlement behaviour in a common coral reef fish. Journal of Animal Ecology 85:903–914. doi: 10.1111/1365-2656.12506

Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR (2014) NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. Molecular Ecology Resources 14:209–214. doi: 10.1111/1755-0998.12157

Dray S, Dufour A-B (2007) The ade4 package: implementing the duality diagram for ecologists. Journal of Statistical Software 22:1–20

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4:359–361. doi: 10.1007/s12686-011-9548-7

Eckman JE (1996) Closing the larval loop: linking larval ecology to the population dynamics of marine benthic invertebrates. Journal of Experimental Marine Biology and Ecology 200:207–237. doi: 10.1016/S0022-0981(96)02644-5

Eldon B, Riquet F, Yearsley J, Jollivet D, Broquet T (2016) Current hypotheses to explain genetic chaos under the sea. Current Zoology 62:551–566

Ellegren H (2014) Genome sequencing and population genomics in non-model organisms. Trends in Ecology & Evolution 29:51–63. doi: 10.1016/j.tree.2013.09.008

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Flowers JM, Schroeter SC, Burton RS (2002) The recruitment sweepstakes has many winners: genetic evidence from the sea urchin *Strongylocentrotus purpuratus*. Evolution 56:1445–1453

Forcucci D, Butler IV MJ, Hunt JH (1994) Population dynamics of juvenile Caribbean spoiny lobster, *Panulirus argus*, in Florida Bay, Florida. Bulletin of Marine Science 54:805–818

Fraser DJ, Debes PV, Bernatchez L, Hutchings JA (2014) Population size, habitat fragmentation, and the nature of adaptive variation in a stream fish. Proceedings of the Royal Society of London B: Biological Sciences 281:20140370–20140370. doi: 10.1098/rspb.2014.0370

Fu C, Fanning LP (2004) Spatial considerations in the management of Atlantic cod off Nova Scotia, Canada. North American Journal of Fisheries Management 24:775–784. doi: 10.1577/M03-134.1

Funes-Rodríguez R, Ruíz-Chavarría JA, González-Armas R, Durazo R, Guzmán-del Proó SA (2015) Influence of hydrographic conditions on the distribution of spiny lobster larvae off the West coast of Baja California. Transactions of the American Fisheries Society 144:1192–1205. doi: 10.1080/00028487.2015.1083474

Gaines S, Roughgarden J (1985) Larval settlement rate: a leading determinant of structure in an ecological community of the marine intertidal zone. Proceedings of the National Academy of Sciences 82:3707–3711

Gardner C, Frusher SD, Kennedy RB, Cawthorn A (2001) Relationship between settlement of southern rock lobster pueruli, *Jasus edwardsii*, and recruitment to the fishery in Tasmania, Australia. Marine and Freshwater Research 52:1271–1275

Gardner C, Hartmann K, Punt AE, Hoshino E (2015a) Fewer eggs from larger size limits: counterintuitive outcomes in a spatially heterogeneous fishery. ICES Journal of Marine Science: Journal du Conseil 72:i252–i259. doi: 10.1093/icesjms/fsu165

Gardner C, Hartmann K, Punt AE, Jennings S (2015b) In pursuit of maximum economic yield in an ITQ managed lobster fishery. Fisheries Research 161:285–292. doi: 10.1016/j.fishres.2014.08.015

George RW (1997) Tectonic plate movements and the evolution of *Jasus* and *Panulirus* spiny lobsters (Palinuridae). Marine and Freshwater Research 48:1121. doi: 10.1071/MF97202

George RW (2005) Review: evolution of life cycles, including migration, in spiny lobsters (Palinuridae). New Zealand Journal of Marine and Freshwater Research 39:503–514. doi: 10.1080/00288330.2005.9517329

Giles EC, Saenz-Agudelo P, Hussey NE, Ravasi T, Berumen ML (2015) Exploring seascape genetics and kinship in the reef sponge *Stylissa carteri* in the Red Sea. Ecology and Evolution 5:2487–2502. doi: 10.1002/ece3.1511

Gimenez L (2006) Phenotypic links in complex life cycles: conclusions from studies with decapod crustaceans. Integrative and Comparative Biology 46:615–622. doi: 10.1093/icb/icl010

Godwin RM, Frusher S, Montgomery SS, Ovenden J (2011) Telomere length analysis in crustacean species: *Metapenaeus macleayi*, *Sagmariasus verreauxi*, and *Jasus edwardsii*. ICES Journal of Marine Science: Journal du Conseil 68:2053–2058. doi: 10.1093/icesjms/fsr144

González-Ortegón E, Giménez L (2014) Environmentally mediated phenotypic links and performance in larvae of a marine invertebrate. Marine Ecology Progress Series 502:185–195. doi: 10.3354/meps10708

Goudet J (2005) Hierfstat, a package for R to compute and test hierarchical F-statistics. Molecular Ecology Notes 5:184–186

Green BS (2008) Maternal effects in fish populations. Advances in Marine Biology 54:1–105. doi: 10.1016/S0065-2881(08)00001-1

Green BS, Gardner C, Kennedy RB (2009) Generalised linear modelling of fecundity at length in southern rock lobsters, *Jasus edwardsii*. Marine Biology 156:1941–1947. doi: 10.1007/s00227-009-1226-x

Green BS, Gardner C, Linnane A, Hawthorne PJ (2010) The good, the bad and the recovery in an assisted migration. PLoS ONE 5:e14160. doi: 10.1371/journal.pone.0014160

Green BS, Gardner C, Hochmuth JD, Linnane A (2014) Environmental effects on fished lobsters and crabs. Reviews in Fish Biology and Fisheries 24:613–638. doi: 10.1007/s11160-014-9350-1

Hart A (2001) Mann-Whitney test is not just a test of medians: differences in spread can be important. British Medical Journal 323:391–393. doi: 10.1136/bmj.323.7309.391

Hastings A, Botsford LW (2006) Persistence of spatial populations depends on returning home. Proceedings of the National Academy of Sciences 103:6067–6072

Hedgecock D (1994) Does variance in reproductive success limit effective population sizes of marine organisms? In: Beaumont A. (ed) Genetics and evolution of aquatic organisms, Chapman & Hall, London, pp 122–134

Hedgecock D, Pudovkin AI (2011) Sweepstakes reproductive success in highly fecund marine fish and shellfish: a review and commentary. Bulletin of Marine Science 87:971–1002. doi: 10.5343/bms.2010.1051

Hendry AP, Day T (2005) Population structure attributable to reproductive time: isolation by time and adaptation by time. Molecular Ecology 14:901–916. doi: 10.1111/j.1365-294X.2005.02480.x

Hendry AP, Berg OK, Quinn TP (1999) Condition dependence and adaptation-by-time: breeding date, life history, and energy allocation within a population of Salmon. Oikos 85:499. doi: 10.2307/3546699

Henning F, Lee HJ, Franchini P, Meyer A (2014) Genetic mapping of horizontal stripes in Lake Victoria cichlid fishes: benefits and pitfalls of using RAD markers for dense linkage mapping. Molecular Ecology 23:5224–5240

Hinojosa IA, Gardner C, Green BS, Jeffs A, Leon R, Linnane A (2017) Differing environmental drivers of settlement across the range of southern rock lobster (*Jasus edwardsii*) suggest resilience of the fishery to climate change. Fisheries Oceanography 26:49–64. doi: 10.1111/fog.12185

Holt RD, Gaines MS (1992) Analysis of adaptation in heterogeneous landscapes: implications for the evolution of fundamental niches. Evolutionary Ecology 6:433–447

Iacchei M, Ben-Horin T, Selkoe KA, Bird CE, García-Rodríguez FJ, Toonen RJ (2013) Combined analyses of kinship and F_{ST} suggest potential drivers of chaotic genetic patchiness in high gene-flow populations. Molecular Ecology 22:3476–3494. doi: 10.1111/mec.12341

Jacobsen MW, Pujolar JM, Bernatchez L, Munch K, Jian J, Niu Y, Hansen MM (2014) Genomic footprints of speciation in Atlantic eels (*Anguilla anguilla* and *A. rostrata*). Molecular Ecology 23:4785–4798. doi: 10.1111/mec.12896

Je Lee H, Boulding EG (2009) Spatial and temporal population genetic structure of four northeastern Pacific littorinid gastropods: the effect of mode of larval development on variation at one mitochondrial and two nuclear DNA markers. Molecular Ecology 18:2165–2184. doi: 10.1111/j.1365-294X.2009.04169.x

Jeffs AG, James P (2001) Sea-cage culture of the spiny lobster *Jasus edwardsii* in New Zealand. Marine and Freshwater Research 52:1419–1424

Jeffs AG, Chiswell SM, Booth JD (2001) Distribution and condition of pueruli of the spiny lobster *Jasus edwardsii* offshore from north-east New Zealand. Marine and Freshwater Research 52:1211–1216

Jeffs AG, Nichols PD, Mooney BD, Phillips KL, Phleger CF (2004) Identifying potential prey of the pelagic larvae of the spiny lobster *Jasus edwardsii* using signature lipids. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology 137:487–507. doi: 10.1016/j.cbpc.2004.02.003

Jeffs AJ, Gardner C, Cockcroft AC (2013) *Jasus* and *Sagmariasus* species. In: Phillips B (ed) Lobsters: Biology, management, aquaculture and fisheries, Blackwell Scientific Publications, Oxford, pp 259–288

Johannesson K, Johannesson B, Lundgren U (1995) Strong natural selection causes microscale allozyme variation in a marine snail. Proceedings of the National Academy of Sciences 92:2602–2606

Johnson GC, Karajah MT, Mayo K, Armenta TC, Blumstein DT (2017) The bigger they are the better they taste: size predicts predation risk and anti-predator behavior in giant clams. Journal of Zoology 301:102–107. doi: 10.1111/jzo.12401

Johnson MS, Black R (1982) Chaotic genetic patchiness in an intertidal limpet, *Siphonaria* sp. Marine Biology 70:157–164

Johnson MS, Wernham J (1999) Temporal variation of recruits as a basis of ephemeral genetic heterogeneity in the western rock lobster *Panulirus cygnus*. Marine Biology 135:133–139

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405. doi: 10.1093/bioinformatics/btn129

Jombart T, Ahmed I (2011) adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27:3070–3071

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC genetics 11:1

Kalinowski S, Manlove K, Taper M (2008) ONCOR: software for genetic stock identification. Department of Ecology, Montana State University, USA.

Kamvar ZN, Tabima JF, Grünwald NJ (2014) Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2:e281. doi: 10.7717/peerj.281

Karlin S, Altschul SF (1990) Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. Proceedings of the National Academy of Sciences 87:2264–2268

Kennington W, Berry O, Groth D, Johnson M, Melville-Smith R (2013a) Spatial scales of genetic patchiness in the western rock lobster Panulirus cygnus. Marine Ecology Progress Series 486:213–221. doi: 10.3354/meps10360

Kennington WJ, Melville-Smith R, Berry O (2013b) Genetics of the wild and captive lobster populations. In: Phillips B (ed) Lobsters: biology, management, aquaculture and fisheries, Blackwell Publishing, Oxford, pp 36–63

Kennington WJ, Cadee SA, Berry O, Groth DM, Johnson MS, Melville-Smith R (2013c) Maintenance of genetic variation and panmixia in the commercially exploited western rock lobster (*Panulirus cygnus*). Conservation Genetics 14:115–124. doi: 10.1007/s10592-012-0433-z

Kerr LA, Cadrin SX, Kovach AI (2014) Consequences of a mismatch between biological and management units on our perception of Atlantic cod off New England. ICES Journal of Marine Science: Journal du Conseil 71:1366–1381. doi: 10.1093/icesjms/fsu113

Kilada R, Sainte-Marie B, Rochette R, Davis N, Vanier C, Campana S, Gillanders B (2012) Direct determination of age in shrimps, crabs, and lobsters. Canadian Journal of Fisheries and Aquatic Sciences 69:1728–1733. doi: 10.1139/cjfas-2012-0254

Kunisch M, Anger K (1984) Variation in development and growth rates of larval and juvenile spider crabs *Hyas araneus* reared in the laboratory. Marine Ecology Progress Series 15:293–301

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nature Methods 9:357–359. doi: 10.1038/nmeth.1923

Larraín MA, Díaz NF, Lamas C, Uribe C, Araneda C (2014) Traceability of mussel (*Mytilus chilensis*) in southern Chile using microsatellite molecular markers and assignment algorithms. Exploratory survey. Food Research International 62:104–110. doi: 10.1016/j.foodres.2014.02.016

Leland JC, Bucher DJ, Coughran J (2015) Direct age determination of a subtropical freshwater crayfish (Redclaw, *Cherax quadricarinatus*) using ossicular growth marks. PLoS ONE 10:e0134966. doi: 10.1371/journal.pone.0134966

Lennon NJ, Lintner RE, Anderson S, Alvarez P, Barry A, Brockman W, Daza R, Erlich RL, Giannoukos G, Green L (2010) A scalable, fully automated process for construction of sequence-ready barcoded libraries for 454. Genome Biology 11:R15

Lesser JHR (1978) Phyllosoma larvae of *Jasus edwardsii* (Hutton) (Crustacea: Decapoda: Palinuridae) and their distribution off the east coast of the North Island, New Zealand. New Zealand Journal of Marine and Freshwater Research 12:357–370. doi: 10.1080/00288330.1978.9515763

Li G, Hedgecock D (1998) Genetic heterogeneity, detected by PCR-SSCP, among samples of larval Pacific oysters (*Crassostrea gigas*) supports the hypothesis of large variance in reproductive success. Canadian Journal of Fisheries and Aquatic Sciences 55:1025–1033

Li L, Stoeckert Jr. CJ, Roos DS (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Research 13:2178–2189

Linck EB, Battey CJ (2017) Minor allele frequency thresholds strongly affect population structure inference with genomic datasets. bioRxiv 188623

Linnane A, Gardner C, Hobday D, Punt A, McGarvey R, Feenstra J, Matthews J, Green B (2010a) Evidence of large-scale spatial declines in recruitment patterns of southern rock lobster *Jasus edwardsii*, across south-eastern Australia. Fisheries Research 105:163–171. doi: 10.1016/j.fishres.2010.04.001

Linnane A, James C, Middleton J, Hawthorne P, Hoare M (2010b) Impact of wind stress anomalies on the seasonal pattern of southern rock lobster (*Jasus edwardsii*) settlement in South Australia: wind stress and puerulus settlement. Fisheries Oceanography 19:290–300. doi: 10.1111/j.1365-2419.2010.00544.x

Linnane A, Sloan S, McGarvey R, Ward T (2010c) Impacts of unconstrained effort: lessons from a rock lobster (*Jasus edwardsii*) fishery decline in the northern zone management region of South Australia. Marine Policy 34:844–850. doi: 10.1016/j.marpol.2010.01.006

Linnane A, McGarvey R, Gardner C, Walker TI, Matthews J, Green B, Punt AE (2014a) Large-scale patterns in puerulus settlement and links to fishery recruitment in the southern rock lobster (*Jasus edwardsii*), across south-eastern Australia. ICES Journal of Marine Science: Journal du Conseil 71:528–536. doi: 10.1093/icesjms/fst176

Linnane A, McGarvey R, Feenstra J, Hawthorne P (2014b) Southern rock lobster (*Jasus edwardsii*) fishery status report 2013/2014. South Australian Research and Development Institute, Adelaide

Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). American Journal of Botany 82:1420–1425

Luikart G, Cornuet J-M (1998) Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conservation Biology 12:228–237

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. Nature Reviews Genetics 4:981–994. doi: 10.1038/nrg1226

Marshall DJ, Uller T (2007) When is a maternal effect adaptive? Oikos 116:1957–1963. doi: 10.1111/j.2007.0030-1299.16203.x

Marshall DJ, Bolton TF, Keough MJ (2003) Offspring size affects the post-metamorphic performance of a colonial marine invertebrate. Ecology 84:3131–3137

Marshall DJ, Cook CN, Emlet RB (2006) Offspring size effects mediate competitive interactions in a colonial marine invertebrate. Ecology 87:214–225

Marshall DJ, Monro K, Bode M, Keough MJ, Swearer S (2010) Phenotype–environment mismatches reduce connectivity in the sea. Ecology Letters 13:128–140. doi: 10.1111/j.1461-0248.2009.01408.x

Marta-Almeida M, Dubert J, Peliz Á, Queiroga H (2006) Influence of vertical migration pattern on retention of crab larvae in a seasonal upwelling system. Marine Ecology Progress Series 307:1–19

Martinsohn JT, Ogden R, Consortium F (2009) FishPopTrace—developing SNP-based population genetic assignment methods to investigate illegal fishing. Forensic Science International: Genetics Supplement Series 2:294–296

Marx JM, Herrnkind WF (1985) Macroalgae (Rhodophyta: Laurencia spp.) as habitat for young juvenile spiny lobsters, Panulirus argus. Bulletin of Marine Science 36:423–431

Mastretta-Yanes A, Arrigo N, Alvarez N, Jorgensen TH, Piñero D, Emerson BC (2015) Restriction site-associated DNA sequencing, genotyping error estimation and *de novo* assembly optimization for population genetic inference. Molecular Ecology Resources 15:28–41. doi: 10.1111/1755-0998.12291

McGarvey R, Ferguson GJ, Prescott JH (1999) Spatial variation in mean growth rates at size of southern rock lobster, *Jasus edwardsii*, in South Australian waters. Marine and Freshwater Research 50:333. doi: 10.1071/MF97172

McGarvey R, Matthews JM, Feenstra JE, Punt AE, Linnane A (2016) Using bioeconomic modeling to improve a harvest strategy for a quota-based lobster fishery. Fisheries Research 183:549–558. doi: 10.1016/j.fishres.2016.05.005

McGrath LL, Vollmer SV, Kaluziak ST, Ayers J (2016) De novo transcriptome assembly for the lobster *Homarus americanus* and characterization of differential gene expression across nervous system tissues. BMC Genomics. doi: 10.1186/s12864-016-2373-3

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research 20:1297–1303. doi: 10.1101/gr.107524.110

McKoy JL, Leachman A (1982) Aggregations of ovigerous female rock lobsters, *Jasus edwardsii* (Decapoda: Palinuridae). New Zealand Journal of Marine and Freshwater Research 16:141–146. doi: 10.1080/00288330.1982.9515957

Merchant S, Wood DE, Salzberg SL (2014) Unexpected cross-species contamination in genome sequencing projects. PeerJ 2:e675

Milano I, Babbucci M, Cariani A, Atanassova M, Bekkevold D, Carvalho GR, Espiñeira M, Fiorentino F, Garofalo G, Geffen AJ, Hansen JH, Helyar SJ, Nielsen EE, Ogden R, Patarnello T, Stagioni M, FishPopTrace Consortium, Tinti F, Bargelloni L (2014) Outlier SNP markers reveal fine-scale genetic structuring across European hake populations (*Merluccius merluccius*). Molecular Ecology 23:118–135. doi: 10.1111/mec.12568

Moland E, Moland Olsen E, Stenseth N (2010) Maternal influences on offspring size variation and viability in wild European lobster *Homarus gammarus*. Marine Ecology Progress Series 400:165–173. doi: 10.3354/meps08397

Morgan EMJ, Green BS, Murphy NP, Strugnell JM (2013) Investigation of genetic structure between deep and shallow populations of the Southern rock lobster, *Jasus edwardsii* in Tasmania, Australia. PLoS ONE 8:e77978. doi: 10.1371/journal.pone.0077978

Morin PA, Luikart G, Wayne RK, the SNP workshop group (2004) SNPs in ecology, evolution and conservation. Trends in Ecology & Evolution 19:208–216. doi: 10.1016/j.tree.2004.01.009

Nicolle A, Moitié R, Ogor J, Dumas F, Foveau A, Foucher E, Thiébaut E (2016) Modelling larval dispersal of *Pecten maximus* in the English Channel: a tool for the spatial management of the stocks. ICES Journal of Marine Science: Journal du Conseil 74:1812–1825. doi: 10.1093/icesjms/fsw207

Nielsen EE, Hemmer-Hansen J, Poulsen NA, Loeschcke V, Moen T, Johansen T, Mittelholzer C, Taranger G-L, Ogden R, Carvalho GR (2009a) Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). BMC Evolutionary Biology 9:276. doi: 10.1186/1471-2148-9-276

Nielsen EE, Hemmer-Hansen J, Larsen PF, Bekkevold D (2009b) Population genomics of marine fishes: identifying adaptive variation in space and time. Molecular Ecology 18:3128–3150. doi: 10.1111/j.1365-294X.2009.04272.x

Nielsen EE, Cariani A, Aoidh EM, Maes GE, Milano I, Ogden R, Taylor M, Hemmer-Hansen J, Babbucci M, Bargelloni L, Bekkevold D, Diopere E, Grenfell L, Helyar S, Limborg MT, Martinsohn JT, McEwing R, Panitz F, Patarnello T, Tinti F, Van Houdt JKJ, Volckaert FAM, Waples RS, Albin JEJ, Vieites Baptista JM, Barmintsev V, Bautista JM, Bendixen C, Bergé J-P, Blohm D, Cardazzo B, Diez A, Espiñeira M, Geffen AJ, Gonzalez E, González-Lavín N, Guarniero I, Jeráme M, Kochzius M, Krey G, Mouchel O, Negrisolo E, Piccinetti C, Puyet A, Rastorguev S, Smith JP, Trentini M, Verrez-Bagnis V, Volkov A, Zanzi A, Carvalho GR (2012a) Gene-associated markers provide tools for tackling illegal fishing and false eco-certification. Nature Communications 3:851. doi: 10.1038/ncomms1845

Nielsen R, Korneliussen T, Albrechtsen A, Li Y, Wang J (2012b) SNP calling, genotype calling, and sample allele frequency estimation from new-generation sequencing data. PLoS ONE 7:e37558. doi: 10.1371/journal.pone.0037558

Ogden R (2008) Fisheries forensics: the use of DNA tools for improving compliance, traceability and enforcement in the fishing industry. Fish and Fisheries 9:462–472

Ogden R, Linacre A (2015) Wildlife forensic science: a review of genetic geographic origin assignment. Forensic Science International: Genetics 18:152–159. doi: 10.1016/j.fsigen.2015.02.008

Ovenden JR, Brasher DJ, White RWG (1992) Mitochondrial DNA analyses of the red rock lobster *Jasus edwardsii* supports an apparent absence of population subdivision throughout Australasia. Marine Biology 112:319–326

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Molecular Ecology 13:55–65

Palma AT, Wahle RA, Steneck RS (1998) Different early post-settlement strategies between American lobsters *Homarus americanus* and rock crabs *Cancer irroratus* in the Gulf of Maine. Marine Ecology Progress Series 162:215–225

Palumbi SR (1994) Genetic divergence, reproductive isolation, and marine speciation. Annual Review of Ecology and Systematics 547–572

Paradis E (2010) pegas: an R package for population genetics with an integrated–modular approach. Bioinformatics 26:419–420

Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–290. doi: 10.1093/bioinformatics/btg412

Parida M, Sannarangaiah S, Dash PK, Rao PVL, Morita K (2008) Loop mediated isothermal amplification (LAMP): a new generation of innovative gene amplification technique; perspectives in clinical diagnosis of infectious diseases. Reviews in Medical Virology 18:407–421

Pechenik JA (2006) Larval experience and latent effects - metamorphosis is not a new beginning. Integrative and Comparative Biology 46:323–333. doi: 10.1093/icb/icj028

Pechenik JA, Tyrell AS (2015) Larval diet alters larval growth rates and post-metamorphic performance in the marine gastropod *Crepidula fornicata*. Marine Biology 162:1597–1610. doi: 10.1007/s00227-015-2696-7

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS ONE 7:e37135. doi: 10.1371/journal.pone.0037135

Phillips BF, McWilliam PS (1986) The pelagic phase of spiny lobster development. Canadian Journal of Fisheries and Aquatic Sciences 43:2153–2163

Piry S (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. Journal of Heredity 95:536–539. doi: 10.1093/jhered/esh074

Plagányi ÉE, McGarvey R, Gardner C, Caputi N, Dennis D, de Lestang S, Hartmann K, Liggins G, Linnane A, Ingrid E, Arlidge B, Green B, Villanueva C (2017) Overview, opportunities and outlook for Australian spiny lobster fisheries. Reviews in Fish Biology and Fisheries. doi: 10.1007/s11160-017-9493-y

Powell SL, Green BS, Leon R, Gardner C (2016) Timing of reproduction in Southern rock lobster, *Jasus edwardsii*, around Tasmania. University of Tasmania, Hobart

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Punt AE, Kennedy RB, Frusher SD (1997) Estimating the size-transition matrix for Tasmanian rock lobster, *Jasus edwardsii*. Marine and Freshwater Research 48:981. doi: 10.1071/MF97017

R Core Team (2016) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Raimondi PT, Keough MJ (1990) Behavioural variability in marine larvae. Australian Journal of Ecology 15:427–437

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. Proceedings of the National Academy of Sciences 94:9197–9201

Reich DA, DeAlteris JT (2009) A simulation study of the effects of spatially complex population structure for Gulf of Maine Atlantic cod. North American Journal of Fisheries Management 29:116–126. doi: 10.1577/M07-092.1

Rey F, Silva Neto G, Brandão C, Ramos D, Silva B, Rosa R, Queiroga H, Calado R (2016) Contrasting oceanographic conditions during larval development influence the benthic performance of a marine invertebrate with a bi-phasic life cycle. Marine Ecology Progress Series 546:135–146. doi: 10.3354/meps11629

Rimmer DW (1980) Spatial and temporal distribution of early-stage phyllosoma of westren rock lobster, *Panulirus cygnus*. Australian Journal of Marine and Freshwater Research 31:485–497. doi: 10.1071/MF9800485

Rodriguez SR, Ojeda FP, Inestrosa NC (1993) Settlement of benthic marine invertebrates. Marine Ecology Progress Series 193–207

Rognes T, Flouri T, Nichols B, Quince C, Mahé F (2016) VSEARCH: a versatile open source tool for metagenomics. PeerJ 4:e2584. doi: 10.7717/peerj.2584

Rosenberg NA, Li LM, Ward R, Pritchard JK (2003) Informativeness of genetic markers for inference of ancestry. The American Journal of Human Genetics 73:1402–1422. doi: 10.1086/380416

Ryman N, Palm S (2006) POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. Molecular Ecology Notes 6: 600-602

Saenz-Agudelo P, Dibattista JD, Piatek MJ, Gaither MR, Harrison HB, Nanninga GB, Berumen ML (2015) Seascape genetics along environmental gradients in the Arabian Peninsula: insights from ddRAD sequencing of anemonefishes. Molecular Ecology 24:6241–6255

Sale PF (1990) Recruitment of marine species: is the bandwagon rolling in the right direction? Trends in Ecology & Evolution 5:25–27

Sanford E, Kelly MW (2011) Local adaptation in marine invertebrates. Annual Review of Marine Science 3:509–535. doi: 10.1146/annurev-marine-120709-142756

Sato T, Hamano K, Sugaya T, Dan S (2017) Effects of maternal influences and timing of spawning on intraspecific variations in larval qualities of the Kuruma prawn *Marsupenaeus japonicus*. Marine Biology. doi: 10.1007/s00227-017-3118-9

Schmid M, Guillaume F (2017) The role of phenotypic plasticity on population differentiation. Heredity 119:214–225. doi: 10.1038/hdy.2017.36

Shanks AL (2009) Pelagic larval duration and dispersal distance revisited. The Biological Bulletin 216:373–385. doi: 10.1086/BBLv216n3p373

Shanks AL, Grantham BA, Carr MH (2003) Propagule dispersal distance and the size and spacing of marine reserves. Ecological Applications S159–S169

Shima JS, Swearer SE (2016) Evidence and population consequences of shared larval dispersal histories in a marine fish. Ecology 97:25–31

Shirley SM, Shirley TC, Rice SD (1987) Latitudinal variation in the Dungeness crab, *Cancer magister*: zoeal morphology explained by incubation temperature. Marine Biology 95:371–376

Slatkin M (2004) Seeing ghosts: the effect of unsampled populations on migration rates estimated for sampled populations. Molecular Ecology 14:67–73. doi: 10.1111/j.1365-294X.2004.02393.x

Smith CT, Seeb LW (2008) Number of alleles as a predictor of the relative assignment accuracy of short tandem repeat (STR) and single-nucleotide-polymorphism (SNP) baselines for Chum Salmon. Transactions of the American Fisheries Society 137:751–762. doi: 10.1577/T07-104.1

Smith KN, Herrnkind WF (1992) Predation on early juvenile spiny lobsters Panulirus argus (Latreille): influence of size and shelter. Journal of Experimental Marine Biology and Ecology 157:3–18

Sorenson L, McDowell JR, Knott T, Graves JE (2013) Assignment test method using hypervariable markers for blue marlin (*Makaira nigricans*) stock identification. Conservation Genetics Resources 5:293–297

Sponaugle S, Cowen RK, Shanks AL, Morgan SG, Leigh JM, Pineda J, Boehlert GW, Kingsford MJ, Lindeman KC, Grimmes C, Munro JL (2002) Predicting self-recruitment in marine populations: biophysical correlates and mechanisms. Bulletin of Marine Science 70:341–375

Stolle E, Moritz RFA (2013) RESTseq – efficient benchtop population genomics with RESTriction fragment SEQuencing. PLoS ONE 8:e63960. doi: 10.1371/journal.pone.0063960

Suzuki N, Sato T (2010) Female size as a determinant of larval size, weight, and survival period in the Coconut crab, *Birgus latro*. Journal of Crustacean Biology 30:624–628. doi: 10.1651/10-3279.1

Teske P, Sandoval-Castillo J, van Sebille E, Waters J, Beheregaray L (2015) On-shelf larval retention limits population connectivity in a coastal broadcast spawner. Marine Ecology Progress Series 532:1–12. doi: 10.3354/meps11362

Teske PR, Sandoval-Castillo J, van Sebille E, Waters J, Beheregaray LB (2016) Oceanography promotes self-recruitment in a planktonic larval disperser. Scientific Reports 6:34205. doi: 10.1038/srep34205

Thomas L, Bell JJ (2013) Testing the consistency of connectivity patterns for a widely dispersing marine species. Heredity 345–354. doi: 10.1038/hdy.2013.58

Thompson AP, Hanley JR, Johnson MS (1996) Genetic structure of the western rock lobster, *Panulirus cygnus*, with the benefit of hindsight. Marine and freshwater research 47:889–896

Tomita N, Mori Y, Kanda H, Notomi T (2008) Loop-mediated isothermal amplification (LAMP) of gene sequences and simple visual detection of products. Nature Protocols 3:877–882

Truelove NK, Griffiths S, Ley-Cooper K, Azueta J, Majil I, Box SJ, Behringer DC, Butler MJ, Preziosi RF (2015) Genetic evidence from the spiny lobster fishery supports international cooperation among Central American marine protected areas. Conservation Genetics 16:347–358. doi: 10.1007/s10592-014-0662-4

Truelove NK, Kough AS, Behringer DC, Paris CB, Box SJ, Preziosi RF, Butler MJ (2017) Biophysical connectivity explains population genetic structure in a highly dispersive marine species. Coral Reefs 36:233–244. doi: 10.1007/s00338-016-1516-y

Veliz D, Duchesne P, Bourget E, Bernatchez L (2006) Genetic evidence for kin aggregation in the intertidal acorn barnacle (*Semibalanus balanoides*). Molecular Ecology 15:4193–4202. doi: 10.1111/j.1365-294X.2006.03078.x

Villacorta-Rath C, Ilyushkina I, Strugnell JM, Green BS, Murphy NP, Doyle SR, Hall NE, Robinson AJ, Bell JJ (2016) Outlier SNPs enable food traceability of the southern rock lobster, *Jasus edwardsii*. Marine Biology. doi: 10.1007/s00227-016-3000-1

Wang M, O'Rorke R, Waite AM, Beckley LE, Thompson P, Jeffs AG (2015) Condition of larvae of western rock lobster (*Panulirus cygnus*) in cyclonic and anticyclonic eddies of the Leeuwin Current off Western Australia. Marine and Freshwater Research 66:1158. doi: 10.1071/MF14121

Waples RS, Do C (2010) Linkage disequilibrium estimates of contemporary N_e using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. Evolutionary Applications 3:244–262. doi: 10.1111/j.1752-4571.2009.00104.x

Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Molecular Ecology 15:1419–1439. doi: 10.1111/j.1365-294X.2006.02890.x

Wei K, Wood AR, Gardner JP (2013) Population genetic variation in the New Zealand greenshell mussel: locus-dependent conflicting signals of weak structure and high gene flow balanced against pronounced structure and high self-recruitment. Marine Biology 160:931–949

Whitlock MC, Lotterhos KE (2015) Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of FST. The American Naturalist 186:S24–S36. doi: 10.1086/682949

Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. Genetics 163:1177–1191

Winter DJ (2012) MMOD : an R library for the calculation of population differentiation statistics. Molecular Ecology Resources 12:1158–1160. doi: 10.1111/j.1755-0998.2012.03174.x

Wong EHK, Hanner RH (2008) DNA barcoding detects market substitution in North American seafood. Food Research International 41:828–837

Wood DE, Salzberg SL (2014) Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biology 15:1